

Good and bad in Pakistan

The Pakistan government is investigating whether scientists leaked nuclear technology to Iran. The affair could undermine public confidence in President Musharraf and divert attention from long-overdue reforms to science and education.

Did Pakistani scientists sell nuclear technology to Iran? And if they did, was the government involved? The answers to these questions cannot be known until the government completes a much-publicized investigation into the activities of its main centre for nuclear-weapons research, the Dr A. Q. Khan Research Laboratories. But it seems that the Islamabad government has already made up its mind who to blame.

The investigation has been prompted by an ongoing probe conducted by the International Atomic Energy Agency into Iran's nuclear affairs. Pakistan's government has acknowledged that Iran has named "certain individuals" as a source of technology transfer in information given to the agency. But the line from Islamabad is that any such transfers were never official policy and are more likely to have been the work of individuals motivated by "personal ambition and greed".

A number of the laboratory's civilian and military staff have been questioned, including Abdul Qadeer Khan, the country's former chief nuclear scientist, who was responsible for enriching the uranium used in Pakistan's 1998 nuclear tests.

The government's statements — coming while the investigation is being carried out — suggest that it is preparing to lay the blame on individual scientists, letting itself off the hook. This doesn't give a very favourable impression of justice in the country and is harsh on those who have a right to a fair hearing. All in all, it seems to be an abdication of responsibility by the army, which has effectively been running the lab since it was set up in 1976.

Towards reform

Alarm bells are ringing, and not just outside the country. The government's handling of the investigation is causing widespread concern inside Pakistan — particularly the fact that some of the scientists caught up in the affair have not been seen by their families since being taken for questioning. So far, the government is refusing to debate the issue in parliament, and press restrictions mean that there is very little comment and analysis in the media.

Pakistan's research community did provide a public show of support for Khan in December. He was guest of honour at a government-sponsored science conference and was praised for his achievements by President Pervez Musharraf's minister for research and higher education, Atta-ur-Rahman, who promised: "We will always stand beside you firmly."

The government's handling of the affair is unfortunate in another respect too: it is diverting attention away from what are perhaps the most far-reaching reforms to Pakistan's science for more than 20 years.

Rahman has convinced Musharraf to prise open his war-chest, and a deluge of money and new initiatives has poured out. Consider the following: the ministry of science and technology's annual budget has increased 60-fold compared with 1999; researchers have an opportunity to more than double their earnings if they publish more in peer-reviewed journals; funds have been released for many more PhDs to be trained at home and abroad; plans are well under way for a free digital library for all educational institutions; there is a new

scheme to attract researchers from overseas to work in Pakistan; and the cabinet recently approved a plan to use science and technology to catalyse a doubling in the country's rate of growth of its gross domestic product before 2020.

Musharraf is inspired partly by Turkey's reforming general Kemal Atatürk, who moved Turkey Westwards following the demise of the Ottoman empire in 1924. Musharraf, a diplomat's son who grew up in 1950s Turkey, sees himself as a Pakistani Kemalist, but with even more ambitious plans to reconcile Islam with modernity. He wants to write the first page of a knowledge renaissance across the Muslim world. Key to this is a plan to establish a US\$1-billion fund for research and development that will be available for the world's 57 countries that have mainly Muslim populations.

Democratic change

All of this is just a bit too giddy for Pakistan's resource-starved research community. Nine years ago, *Nature* sent a correspondent to survey the Pakistani science scene. The science ministry had no minister. Its total annual research spending was just \$7 million and the country produced just two PhDs per university per year. This is probably closer to the country's default research environment and many researchers are worried that Musharraf and Rahman have set standards that will be impossible to match once the general leaves office.

High-quality research needs more than money and good ideas. It needs an environment that supports democracy, the freedom to think, write and publish; an appreciation of the value of multiculturalism and diverse opinions; and the rule of law. Many Muslim countries, it has to be said, do not support such an environment. Pakistan is among the better ones and under Musharraf has taken faltering steps in the right direction.

Building confidence

What is unfortunate is that the government's current handling of the nuclear affair supports the view held by many commentators, inside and outside Pakistan, that the Musharraf regime is in effect little more than an old-style military dictatorship. The apparatus of democracy may be visible — there is a functioning parliament, local councils and a judiciary — but the spirit seems to be missing. There is no real way to challenge decisions made by the executive and no proper access to justice.

It is rare to see a research minister wielding much — if any — influence over a head of state. But Rahman clearly appears to have his leader's confidence. Musharraf recently spoke of him in glowing terms: "He has taken me to a land of fantasies so many times that I now live in — and believe in — the land that he imagines. It is so real and so motivating that all I as a layman can do is to give him total support and back-up without reservation and allow him free vent and financial support for his ideas."

If Rahman has as much influence over his boss as it seems, now is the time for him to make it clear that lasting reforms need more than an injection of money, and that there should be no return to the bad old days. ■





Express delivery
High-resolution
images of the martian
surface reach Earth
p382



Down by law
Japan faces criticism
over ethnic profiling
of suspects
p383



Under attack
Pressure mounts on
US efforts to conserve
predators
p385



Golden shot
German proposal
would sell reserves
to pay for science
p386

Prolific ecologist vows to fight Danish misconduct verdict

Alison Abbott, Munich

Denmark's main research misconduct body has ruled that Anders Pape Møller, one of the world's leading experts in behavioural ecology, is guilty of "scientific dishonesty".

Møller, a population biologist now at the Laboratory of Evolutionary Parasitology run by the CNRS, France's national research agency, at the Pierre and Marie Curie University in Paris, says that he is innocent and is appealing against the ruling.

The verdict was published last November on the website of the Danish Committees on Scientific Dishonesty (DCSD), but only circulated internationally after an English translation appeared last week on the website of the University of Copenhagen's Zoological Institute. It has rocked the tight-knit community of which Møller is part — he has been prolific in many areas of behaviour, evolution and ecology, and many of his findings are incorporated into standard textbooks.

"It will be a matter of no small consequence if a question mark appears over any part of Møller's work," says Chris Barnard, a behavioural ecologist at the University of Nottingham, UK, and president of the Association for the Study of Animal Behaviour (ASAB). Barnard says that he is waiting to see the outcome of Møller's appeal before deciding whether the ASAB should take any action. The society co-publishes the journal *Animal Behaviour*, which has carried several of Møller's papers.

The DCSD investigated an allegation brought in 2001 by Jørgen Rabøl, a former colleague at the Copenhagen institute, where Møller worked between 1994 and 1996. It concerns a paper testing whether herbivore activity causes asymmetry of new leaf growth in oak trees¹. This is part of a large body of work by Møller suggesting that symmetry is an evolutionarily important signal of genetic quality — used by animals in mate choice, for example.

The paper was retracted by Møller in March 2001, with a note saying that "the measurements and analyses behind the data



Under question: animal behaviour, particularly of barn swallows, has been the focus of ecologist Anders Pape Møller's work.

... were flawed and misinterpreted"². But Rabøl claimed that the wording of the retraction wrongly implied that a technician working with Møller had been to blame. He suspected that the technician's data had not been used in the paper, and filed a complaint to the DCSD.

Data dispute

The DCSD set up a three-person committee to investigate. Møller was unable to provide original data; instead he offered a transformed data set that he said corresponded to three tables in the paper. But the committee concluded that this could not have been based on authentic measurements — in part because many data points were identical. "There are very strong indications that it must, at least in part, be fabricated," says the committee's report.

The report also notes that the original measurements submitted by the technician did not agree with the results reported in the paper, and so could not have been used. It says that Møller's reference in his retraction to flawed measurements is "hardly credible".

Even before the verdict Møller had filed a complaint with the Danish research ministry, claiming that one member of the investigating committee had a conflict of interest. Since the posting of the verdict in English, Møller has circulated this complaint by e-mail to international colleagues. This note alleges that Rabøl had personal grudges against him, and a history of making mis-

conduct allegations. It also claims that the technician involved was an alcoholic. Bo Vest Pedersen, director of the Copenhagen institute, disputes this last point.

Møller has been taken to task before by colleagues. In 1998, evolutionary biologist David Houle of Florida State University in Tallahassee criticized a book, *Asymmetry, Developmental Stability and Evolution*, co-authored by Møller. In a review in *Evolution*³, Houle wrote that Møller had plagiarized a 200-word section from another researcher's unpublished manuscript that he was reviewing.

A year later, evolutionary ecologist Richard Palmer of the University of Alberta in Edmonton, Canada, wrote in *The American Naturalist* that Møller may have selectively published data to make his point⁴.

Some of Møller's co-authors are also aggrieved. Andrew Pomiankowski of University College London, who wrote four papers with Møller in the 1990s, says that he stopped working with him, suspecting unreliability. "If his name is on a paper, I will not read it," says Pomiankowski.

But Møller's supporters counter that critics are jealous of Møller's success and his vast publication record — he co-authored more than 100 papers between 2001 and 2003.

Møller says that he expects to be cleared by the ministry, but if not he vows to take his case to the courts or the Danish ombudsman. He notes that the DCSD last year landed in hot water after charging Bjørn Lomborg, author of a controversial book which claimed that environmentalists have exaggerated the threat to our planet, with dishonesty. That verdict was challenged by the Danish government (see *Nature* 427, 7; 2004).

Jean Clobert, former director of Møller's Paris lab, says that he is taking advice from the CNRS about how to respond. A second, CNRS-led, inquiry cannot be ruled out, he says.

Additional reporting by Rex Dalton.

1. Møller, A. P. & de Lope, F. *Oikos* **82**, 246–252 (1998).
2. Møller, A. P. & de Lope, F. *Oikos* **92**, 558 (2001).
3. Houle, D. *Evolution* **52**, 1872–1876 (1998).
4. Palmer, A. R. *Am. Nat.* **154**, 220–233 (1999).

Prospects brighten around the red planet

Tony Reichhardt, Washington.

Planetary scientists are revelling in a flood of fresh data from Mars. A second NASA rover landed on the planet's surface on 24 January, the first seems to be on the mend after a computer glitch, and the European Space Agency's (ESA's) Mars Express has begun transmitting spectacular photographs and spectral data from martian orbit.

NASA's Opportunity spacecraft, which came down in a rocky plain known as Meridiani Planum, rolled to a stop inside a shallow, 20-metre-wide crater that partially blocked its view of the horizon. But once the rover's systems are checked out, project engineers predict that it will easily get out of the crater. The terrain at Meridiani is different from any previous landing site on Mars, with darker soil and at least one exposed outcrop that may offer a rare glimpse of martian bedrock.

The two rovers are searching for geological evidence of water, past or present, on Mars. The first spacecraft, Spirit, set down in what scientists believe is a dried-up lake bed, whereas the Meridiani site contains an iron mineral — haematite — that forms in the presence of water on Earth.

Spirit went silent on 20 January just as it was about to begin its first close-up studies of a rock. After a nervous 24 hours without radio contact, controllers at NASA's Jet



The Mars probe Opportunity prepares to climb the crater wall near its landing spot.

Propulsion Laboratory in Pasadena, California, regained command of the spacecraft. They determined the probable cause of the blackout — a problem with onboard computer memory caused by handling too many data files. Although it may take a couple of weeks to debug the software and check out onboard systems, project managers are optimistic that Spirit will be able to resume exploration. Project scientist Steven Squyres of Cornell University, Ithaca, New York, said the team had anticipated before the mission that such glitches could interrupt science operations as often as one day out of three.

ESA's Mars Express, meanwhile, is nearing the end of its checkout phase in orbit, with all instruments working perfectly, according to project scientist Agustin Chicarro. ESA researchers last week unveiled the first pictures

from the orbiter's main camera (see below), as well as data from other instruments. A combined camera and infrared spectrometer called OMEGA has produced the first significant scientific results from the mission: direct detection of water ice and carbon dioxide ice in the planet's south polar cap.

Although some scientists have downplayed the significance of this result, past telescopes and satellites saw only atmospheric water vapour or hydrogen on the martian surface. The direct detection of water molecules in the polar caps is "obviously not a surprise — it's what everybody expected", says Chicarro. But confirming it is a significant step, he says.

Unlike the pictures from the twin NASA rovers, data and images from Mars Express will at first belong to the ESA scientists who built the instruments. After six months, they will be placed in a public archive — an arrangement similar to that for the Mars Observer Camera now orbiting the planet on NASA's Mars Global Surveyor.

Malin Space Science Systems of San Diego, California, which built that camera, last week released a picture showing the Spirit rover, its parachute and other cast-off hardware as bright dots on the martian surface — the clearest image ever taken of a planetary lander from orbit.

Photo express

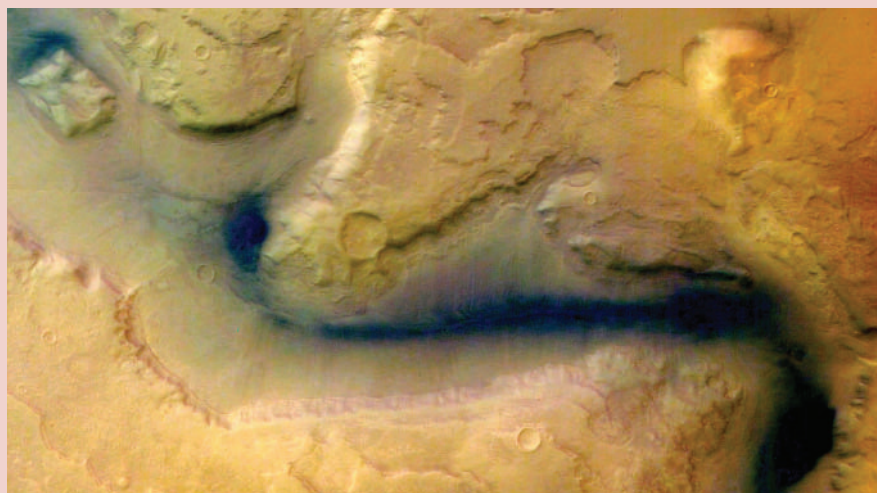
Quirin Schiermeier, Munich

Spectacular pictures from the European Space Agency's Mars Express mission are being beamed back to Earth each day now. Their beauty and clarity are thrilling geologists — and laying the foundations for the richest database ever gathered of any planet, even our own.

The pictures are the product of the High Resolution Stereo Camera (HRSC). The €70-million (US\$87-million) instrument was originally developed by scientists at the German Space Agency for Russia's failed Mars '96 mission and has been updated for Mars Express.

Digital processing helps to create the stunning three-dimensional views, but instrument builders say that the pictures are 'real'. Like a normal camera, the HRSC takes photographs through a single lens, which is pointed at the planet's surface each time Mars Express's elliptical orbit reaches its lowest altitude, about 250 kilometres.

Behind the lens, however, is a complex system of parallel sensors, which are sensitive to red, green, blue and near-infrared light. By shooting points on the planet's surface from three different perspectives — forwards, downwards and backwards — the camera collects information that is then processed back on Earth to create a digital model of the terrain with a spatial resolution of 10 metres per pixel or less.



Past evidence: Reull Vallis, a channel once formed by flowing water, as seen from Mars Express.

As of 23 January, after 10 days of observation, 2 million square kilometres — or about 2% of the Mars surface — had been surveyed. After some 5,000 flyovers, which would take a maximum of four Earth years to complete, the topography of Mars would have been surveyed more comprehensively than that of Earth, where military considerations prevent complete topographical surveys from reaching public databases.

"We can clearly make out slopes and valleys that were almost certainly eroded by running

water," says Gerhard Neukum, a planetary scientist at the Free University in Berlin, and principal investigator on the HRSC project. "Moreover, we have found moraines, which proves there must have been glaciers on Mars."

The camera might also be able to spot the large white parachutes of the Beagle 2 lander, which went missing on 25 December last year. But so far the search has been in vain. "There is dust and lack of contrast where we assume it must be," says Neukum.



Trial and error: Japanese police want to link biological evidence to a perpetrator's ethnic background.

Japan's ethnic crime database sparks fears over human rights

David Cyranoski, Tokyo

Ethnic profiles derived from biological material at crime scenes are set to become an integral part of criminal investigations in Japan.

The move has provoked criticism from scientists and human-rights activists who fear that the data will be used to target foreigners unfairly.

In a four-year, ¥153-million (US\$1.4-million) project, forensic scientists at the National Research Institute of Police Science in Tsukuba, north of Tokyo, plan to set up a database that will allow police to get clues about a suspect's ethnic origin and physical appearance.

The Japanese authorities cite the increase in crimes committed by foreigners as justification for the project. Together with research into identifying the origin of falsified passports, the ethnicity indicators will "help in the investigation of crimes involving foreigners, and by leading to a quick resolution of these crimes, effectively reduce their number and decrease social unrest", according to the initiative's outline.

But Makoto Teranaka, head of Amnesty International's Tokyo office, describes the project as a "sign of a moral panic", adding that it compounds Amnesty's long-standing concerns about how Japan's prosecution system attains its stunningly high conviction rate of about 99%. "We want Japan to bring its system up to international standards for fair trial and interrogation," he says.

The database will gather data on ethnicity, blood type, metabolic enzymes, hair and skin pigment proteins, mitochondrial DNA and signs of asymptomatic viral infections that can be used to distinguish ethnicities.

There is nothing new about using biological features, as in DNA profiling, to identify

criminals, but ethnicity data are usually only used in the identification of badly damaged corpses. The Japanese initiative is thought to be the first that will try to attach an ethnic profile to an individual's DNA.

A number of Japanese researchers share Amnesty's concern about the project, although they declined to criticize it openly. One prominent geneticist at a leading Japanese university, contacted by *Nature*, dismissed the plan as "stupid" but refused to comment on it publicly.

And Japanese forensic scientists, whose research is being redirected from corpse identification for the project, are worried. "It gets a little difficult ethically," says one, who didn't want to be identified. "We have to worry about human rights, especially when the tests have some uncertainty built into them." But the researcher also agrees with the authorities that growing crime by foreigners might justify the initiative.

Scientists at the police-science institute insist that the tools will "only be used to narrow in on foreign criminals or victims and not to discriminate against individuals or ethnicities". But critics charge that the initiative fits in with a pattern of actions that scapegoat foreigners for Japan's social and economic problems.

Koichi Hamai, a criminologist at Ryukoku University in Kyoto who worked at the justice ministry for 19 years, worries that the initiative might encourage the police to focus on foreigners. "It could become a dangerous prejudice," he says. The number of foreigners convicted of murder — 41 in 2003 — is actually decreasing, Hamai adds, and theft convictions, although increasing, amounted to only 280 cases out of 4,151. "Foreigners have become a scapegoat in the years since the bubble economy burst," he says. ■

North Korea offers US tour party glimpse of weapons programme

Geoff Brumfiel, Washington

Speeding along a deserted highway towards North Korea's secret city of Yongbyon earlier this month, Siegfried Hecker, former director of the Los Alamos nuclear-weapons laboratory in New Mexico, wasn't sure what to expect.

"I had heard that the North Koreans were less than cooperative, to put it kindly," he says. In the end, Hecker, who was invited to make the semi-official trip by North Korea's nuclear-weapons researchers, came away impressed by their technical expertise, but unsure of the status of their weapons programme.

Hecker, who told a closed Senate hearing on 21 January about his visit and later discussed it with reporters, says he engaged in a surprisingly frank and even humorous discussion with his North Korean counterparts. Ri Hong Sop, the director of the Yongbyon facility, gave Hecker and four other US visitors a guided tour on 8 January and sought to convince them that North Korea had successfully assembled nuclear weapons.

The visitors were shown a cooling pond where 8,000 spent fuel rods had been housed in steel canisters. The North Koreans claimed that all of them had been reprocessed to extract plutonium, but many of the canisters looked to be intact. "I said, 'I can't go back home and report that all the fuel rods are missing, because I didn't see them all empty,'" Hecker recalls. So Sop suggested that he test the claim by choosing a canister to be opened at random. When he opened a random canister and found it empty, he concluded that the fuel rods were indeed missing.

Later, Hecker was allowed to handle a mildly warm metal sample, which the North Koreans claimed was bomb-grade plutonium alloy. Hecker, a metallurgist who knows plutonium and its alloys rather well, asked Sop what he had used to create the alloy. "He said, 'I'm not authorized to tell you, but you know, it's the same stuff that you use,'" Hecker says he is reasonably convinced that what he held could have been weapons-grade plutonium.

But Hecker is less sure that North Korea has turned its plutonium into a bomb. A plutonium-based weapon is more difficult to build than the enriched-uranium device that some analysts assumed North Korea would try to build. But Hecker was impressed by the technical abilities of some of the scientists he met. "The people were very competent, no question about it," he says. ■

Quashed convictions reignite row over British cot deaths

Laura Nelson, London

The British government is set to review the cases of 258 British parents convicted of murdering their own babies in the past decade, after the Court of Appeal threw out the expert testimony used in three such convictions.

A row has erupted over testimony given by retired paediatrician Roy Meadow. In 1977, Meadow, then at St James's University Hospital in Leeds, proposed the condition 'Munchausen syndrome by proxy' (R. Meadow *Lancet* ii, 343–345; 1977), which describes parents who deliberately harm their children to draw attention to themselves. In many court cases he testified on the probability that this syndrome was responsible for sudden infant death syndrome, or SIDS.

Meadow has been pilloried in the British press for his role as expert witness in the three convictions, and is under investigation by the General Medical Council, which governs British medicine. He is due to appear before its professional conduct committee this summer.

Lawyers representing Meadow have advised him not to comment on the charges, he says. But some scientists have privately sprung to his defence. One senior paediatrician contacted by *Nature* said that Meadow "has made important contributions" to the understanding of cot deaths and had conducted himself with "skill and compassion".

Others feel that Meadow has allowed his judgement to be clouded. "Meadow has gone over the top," says Robert Carpenter, who researches SIDS at the London School of Hygiene and Tropical Medicine. "He sees child abuse behind every tree."

Joyce Epstein, director of the London-based charity, the Foundation for the Study of Infant Deaths, called for more thorough investigation into each death. "There should also be an examination of 40 different tissues, including the whole brain," says Carpenter.

Peter Fleming, a child-health expert at the University of Bristol, will publish an article next month in the *British Medical Journal* calling for such detailed investigations to be performed within 24 hours of the baby's death. But doctors are reluctant to enter the ethical and legal minefield of such investigations, making a uniform regime tough to implement. ■



Death dealer: the tsetse-fly *Glossina morsitans* transmits African sleeping sickness.

African labs win major role in tsetse-fly genome project

Declan Butler

An international research consortium has been formed to sequence the genome of the tsetse-fly, the pest that carries disease to millions of people and cattle in Africa.

The tsetse-flies, species of *Glossina*, carry trypanosome parasites that cause human sleeping sickness (African trypanosomiasis) and animal trypanosomiasis, a scourge of African cattle that is estimated to cost farmers US\$4.5 billion a year.

The International *Glossina* Genomics Initiative, which was launched at a meeting in Geneva on 19–20 January, plans to break new ground by involving African scientists from the start. If the initiative is funded, most of the high-throughput sequencing will be done in Europe and the United States, but the databases will be housed at the South African National Bioinformatics Institute (SANBI) at the University of Western Cape, near Cape Town. Much of the genomic and bioinformatics analysis will be carried out in Africa.

African trypanosomiasis infects up to half a million people, and is fatal if untreated. The main therapy available, a 50-year-old drug called melarsoprol, kills up to 10% of those who take it. With no new drugs in view for at least a decade (see *Nature* 424, 10–11; 2003), a sequenced genome could help researchers better understand the interaction of the pathogen with its fly and human hosts and its transmission to humans, says Serap Aksoy, head of Yale School of Public Health's Division of Epidemiology of Microbial Diseases, who attended the meeting. "This is a developing-country problem; it's

very important that developing countries be involved from the beginning," she adds.

Winston Hide, director of SANBI, says that the inclusion of African nations recognizes their efforts to create a workforce skilled in genomics. The South African government has pledged \$40 million over ten years to create a national bioinformatics network. Over the past four years, the World Health Organization has also funded bioinformatics training courses in Africa, Latin America and Asia. "We have trained a cadre of scientists who are from very poor places, but highly qualified," says Hide.

The consortium was formed at a meeting held under the auspices of the United Nations Tropical Disease Research programme. It brings together scientists from more than a dozen sleeping-sickness labs and genome centres, many of which are already sequencing the *Trypanosoma brucei* genome.

The first goal will be to check whether the choice of fly to be sequenced — *Glossina morsitans* — is a valid model for the more important vector of human disease, *G. palpalis*. The latter's genome is full of repeat sequences and weighs in at some 7,000 million bases (Mb), or twice the size of the human genome. That of *G. morsitans*, at 600 Mb, is more manageable.

The costs of this first nine-month phase, which would involve small amounts of sequencing, can be covered by coordinating existing funds among the consortium partners, says Aksoy. The consortium would then seek \$10 million to sequence and annotate the complete genome by 2006. ■

Wildlife attacks hinder conservation efforts

Rex Dalton, San Diego

Incidents such as the untimely death of Mark Jeffrey Reynolds, a 35-year-old cyclist who was disembowelled by a puma in the Whiting Ranch Wilderness Park near Los Angeles on 8 January, don't make it any easier for biologists who are trying to study and conserve predator species in the American West.

Yet efforts to conserve pumas, wolves, grizzly bears and other predator species have continued to gain momentum. Researchers hope that more thorough studies of animal behaviour will reduce the number of accidents and maintain the fragile public support for the conservation of even the fiercest predators.

Only days after Reynolds and another cyclist, Anne Hjelle, who survived, were attacked, researchers at the University of California, Davis, released what they say is the most comprehensive study yet of interactions between pumas and people.

"Ultimately, it is up to people to decide if they want to share the environment with an animal that can kill them," says Walter Boyce, director of the university's Wildlife Health Center and leader of the study. "Attacks are rare, but there is no guarantee of safety."

Boyce and colleagues charted the lives of 20 pumas for 3 years in the Cuyamaca Rancho State Park near San Diego, close to where a birdwatcher was killed by a puma in 1994.



Lethal potential: in the face of recent puma attacks, a wildlife study gives safety recommendations on living with California's growing puma population.

The observations, made using radio-collars that beamed the pumas' locations to a satellite, showed that the animals hunted deer and bighorn sheep, and cached their kills within a few hundred metres of people's homes and a girls' camp. But they found that the pumas were careful to avoid humans, and were most active between dusk and dawn.

Boyce says that his group's findings will help visitors to the pumas' territory make informed decisions about their own safety. Safety recommendations from the study are being widely circulated in California, where puma populations have been growing in protected wilderness.

Pumas are abundant in California, but rarer predators are the subject of contentious

protection battles. For example, the grey wolf, which was wiped out in the United States 75 years ago, has been successfully reintroduced since 1995 in Montana, Idaho and Wyoming.

Conservationists see this reintroduction as one of the great successes of the 1973 Endangered Species Act (ESA). But ranchers, hunters and associated businesses are objecting to the growing population of grey wolves outside the Yellowstone National Park, where the conservation effort started.

The US Fish and Wildlife Service (FWS), which controls the effort, was prepared to relax ESA protection of the wolf in

the three states if they came up with satisfactory plans to conserve the species on their own. Earlier this month, FWS officials approved the plans offered by Montana and Idaho, but rejected that of Wyoming. The plan drawn up by this staunchly conservative state would have allowed widespread killing of the wolves in some areas.

Wyoming officials say that a conservation effort would cost \$700,000 a year, and that the federal government should pay. A similar dispute is expected to arise later this year, when the FWS will try to transfer responsibility for conservation of the grizzly bear to Wyoming and other western states.

► www.vetmed.ucdavis.edu/whc/scp/mnt_lion.htm

NIH acts to quench 'conflict of interest' allegations

Erika Check, Washington

The National Institutes of Health (NIH) looks set to ensure fuller disclosure of dealings that its senior staff have with private companies.

Elias Zerhouni, director of the biomedical research agency, told a congressional hearing on 22 January that the agency would take steps to guard against conflict of interest, or the perception of it, among its senior staff. "I have reached the conclusion that the NIH must make changes that will appropriately restrict current practices," he said.

The hearing was held in response to an investigation by the *Los Angeles Times*, which reported on 7 December last year that some senior NIH scientists were receiving consulting payments from, or held shares in, biotechnology companies that were benefiting from grants or other decisions

that the scientists could have influenced (see *Nature* 426, 739; 2003). The newspaper noted that, under rules established in 1995, the scientists were not required to publicly disclose these interests.

Zerhouni also announced that a panel co-chaired by Bruce Alberts, president of the National Academy of Sciences, and Norman Augustine, chairman of the Lockheed Martin Corporation, would look into the allegations and report within 90 days on steps the agency could take to guard against future conflicts of interest.

Senator Arlen Specter (Republican, Pennsylvania), chair of the Senate appropriations subcommittee that held the hearing, said the allegations raised serious issues for the agency. "I believe there have to be substantial remedial steps taken to ensure that a wall of separation between public duties and private

compensation is maintained," Specter said.

Observers say it is now likely that the US Department of Health and Human Services, of which the NIH is part, will change the rules that currently shield some agency scientists from public scrutiny. On 12 January, health-department lawyer Edgar Swindell sent a letter to the Office of Government Ethics, which oversees ethics regulations for the US federal government. He requested that senior officials at the NIH's 27 centres and institutes be required to disclose their sources of income to the public. At the Senate hearing, an official from the ethics office signalled that she would approve this request.

According to documents released by the health department, Swindell made a similar request six years ago which the ethics office rejected on technical grounds.

Cambridge abandons primate centre plan as security costs rocket

London The University of Cambridge, UK, has scrapped its plans for a new primate facility for neuroscience research. The decision, made as *Nature* went to press, was forced by rising costs due in part to long-running protests by animal-rights activists.

Capital costs for the facility rose from £24 million (US\$43.4 million) to £32 million during a series of planning applications and appeals over the past three years. The government gave the go-ahead last November, but the university pulled out on 26 January. A spokesman said that extra running costs needed to maintain security at the facility contributed to the decision.

"This decision is very unfortunate," says Mark Walport, director of the Wellcome Trust, a medical charity that pledged £22 million towards the project. "But it will not change our position on what we fund."

Germany goes for gold to bolster science funding

Munich Scientists in Germany have a golden chance to benefit from a windfall, following

a proposal to sell up to 600 tonnes of gold reserves as a potential boost to science and innovation.

Government science experts last week suggested that selling the gold could create a fund of several billion euros. The interest, estimated at €250 million (US\$312 million) per year, could then be used for research, they said.

Germany holds the world's second-largest gold reserves, roughly 3,500 tonnes, currently worth around €37 billion. The government has in the past suggested selling gold to fill gaps in public budgets. Germany's central bank, which controls the reserves, has always declined the idea.

But now Ernst Welteke, president of



Gold standard: German science could benefit from the sale of 600 tonnes of gold reserves.

Germany's central bank, has backed the plan. However, using the bank's revenues for a specific purpose, such as research and education, will require a change to its legal statutes, he said.

Science administrators have welcomed the proposal. "The feasibility of selling gold for research purposes needs to be examined very seriously," says Karl Max Einhäupl, the president of the Wissenschaftsrat, Germany's science advisory council.

Monsanto tries to engineer wheat export

Cape Town The biotechnology firm Monsanto is preparing to export genetically modified wheat to South Africa, pending US approval of the transgenic crop.

The company has applied to the South African government to export its genetically engineered wheat to the country. The crop is currently not grown commercially anywhere in the world. If US clearance is not granted, however, the export of the wheat to South Africa would contravene the Cartagena protocol, to which South Africa is a signatory. This stipulates that transgenic crops cannot be exported until they have been cleared for domestic consumption.

Shannon Troughton, a spokeswoman for Monsanto, says the notice is a normal part

Coordinated effort aims to eradicate polio

New Delhi Traditional foes India and Pakistan have joined hands to fight a common enemy — polio. As well as holding coordinated immunization drives on the same days, they will form a joint surveillance team to monitor the border areas for new cases of polio. The programme is due to start on 22 February, when teams from both countries will immunize children under the age of five.

India and Pakistan reached this agreement at a World Health Organization conference in Geneva on 15 January. The six countries where polio is endemic — India, Pakistan, Nigeria, Niger, Afghanistan and Egypt — signed a



declaration pledging to eradicate polio by the end of this year.

of the process of seeking markets once the US government has given its approval. Monsanto will adhere to international treaty requirements if and when it exports the wheat, she adds.

Congress boosts defence and security in 2004 budget

Washington After months of wrangling, the US Congress has finally settled the budget for the fiscal year 2004, which began last October. And researchers outside the defence community are getting

only a modest increase in funding.

The budget includes major gains for the newly formed Department of Homeland Security, whose research budget rose by more than 50% to US\$1 billion. Research and development at the Department of Defense rose by 13% to \$66 billion.

Non-defence research agencies fared less well. The National Institutes of Health received a modest 3.2% increase to \$27 billion, with most of the new money targeted at biodefence projects.

The National Science Foundation received a 4.7% boost to just over \$4 billion,

well below the 15% increase that advocates had called for. The Department of Energy's office of science received a 3.8% increase to \$3.2 billion. The National Institute of Standards and Technology's funding fell by 4% to \$500 million, although its Advanced Technology Program survived a threatened closure.

US medical biodefence agency proposed

Washington The efforts of the US Department of Defense to find new countermeasures against biowarfare are failing, according to a report from the Institute of Medicine.

The report, published on 22 January, calls for the defence department to set up a new medical biodefence agency, which would consolidate existing research programmes, worth about \$320 million annually.

It also notes that the department has not developed any new vaccine for more than ten years. "The committee sees dismal prospects for successful results" from the current Chemical and Biological Defense Program, the report says. Compiled by a committee led by the pharmacologist Leslie Benet of the University of California, San Francisco, it says that the department's programmes are plagued by poor organization.

The reluctant celebrity

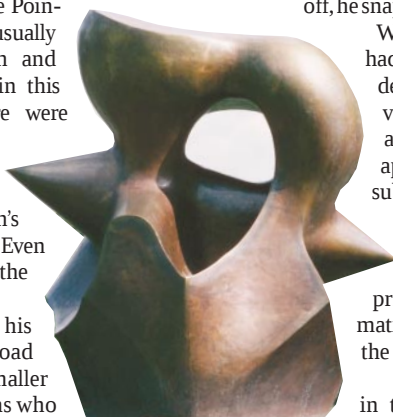
A reclusive Russian claims to have solved a century-old mathematical problem — but his enigmatic personality is adding a fresh dimension to the proof-checking process. Emily Singer reports.

It takes guts for a young Russian to stride into the heart of the US mathematical establishment after claiming to have solved a problem that has baffled the discipline's finest minds for decades. But by the time that Grigory 'Grisha' Perelman, a 30-something recluse, arrived to talk at the Massachusetts Institute of Technology (MIT) in April 2003, it was clear that there was some substance behind his bravura.

Five months before, in a posting on the Internet¹, Perelman claimed to have proved the geometrization conjecture, a theory crucial for understanding three-dimensional surfaces. This would automatically prove the more famous Poincaré conjecture, which has stumped mathematicians for 100 years². The world's top mathematicians had tried to pick holes in his argument — and although their job was far from complete, they already suspected that Perelman was onto something big.

"Every few years, someone claims to have solved the Poincaré conjecture. It is usually a chore to go through and find the problem. But in this case it was clear there were new and brilliant ideas," says Tom Mrowka, a mathematician at MIT who has studied Perelman's work. "It was so original. Even if there was a gap near the end, it was a big deal."

When Perelman gave his talks at MIT, both to a broad audience and to a smaller group of mathematicians who had spent several months studying his work, he had answers for every question



Mathematically this Henry Moore sculpture is the same as a doughnut.



Proving the Poincaré conjecture will offer fresh insight into three-dimensional geometry.

that came up. "It was clear he had thought about all these issues before," says Mrowka. "He'd either point out that it was a trivial question or he'd have an answer."

The media soon got wind of the potential significance of Perelman's achievement. An article about his work appeared in *The New York Times* in mid-April, and when Perelman lectured at the Courant Institute of Mathematical Sciences in New York two weeks later, reporters mingled with mathematicians in the audience. This time, Perelman was much less effusive. He refused to answer reporters' questions or speculate about the implications of his work. When a photographer's flash went off, he snapped: "Don't do that!"

Within weeks Perelman had returned to Russia, evidently annoyed at the unwanted publicity. He ignored a flurry of job offers and apparently has no plans to submit his work to a peer-reviewed journal. He has also shown no interest in a US\$1-million prize that awaits the mathematician who finally proves the Poincaré conjecture.

Why did Perelman react in this way to the acclaim generated by his work? It's hard to say. He talks about little other than mathe-

matics, even with those who count themselves as his friends. Colleagues have only theories for his reclusive behaviour. And e-mails from *Nature*, requesting an interview for this article, went unanswered.

Scratching the surface

Perelman's work focuses on the geometrical properties of three-dimensional surfaces. In mathematical terms, the thin film that makes up a soap bubble is a two-dimensional surface that curves round to enclose a three-dimensional space. Similarly, there is a three-dimensional equivalent of the bubble's surface, called a three-sphere. Understanding such shapes could help mathematicians solve a host of topological problems, and perhaps even describe the shape of the Universe.

Nineteenth-century mathematicians had shown that any closed two-dimensional surface can be described — at a fundamental level — as one of two basic shapes: a sphere or a 'doughnut' with one or more holes. An egg is, in essence, the same as a smooth sphere; a coffee mug is a doughnut. Mathematically, an important characteristic separates spheres from doughnuts. A loop stretched around a sphere can always be shrunk down to a point — just as an elastic band round an egg can be pulled tight to a single point without losing contact with the egg's surface. This is not possible for a doughnut — a loop passing through the doughnut's hole cannot be pulled to a point without cutting through the doughnut

ously set out to prove Thurston's conjecture using the Ricci flow, a systematic procedure that smooths an object's surface into a simpler — homogenous — shape by spreading its curvature. But this isn't always easy. Some parts of the surface may transform faster than others, resulting in a 'lumpy' shape. These problem points, called singularities, prevented Hamilton from succeeding.

"You need to control the way singularities form," explains Jim Carlson, president of the Clay Mathematics Institute in Cambridge, Massachusetts. "Perelman found new inequalities that allow you to do that."

If it is correct, Perelman's work will have provided a proof for both the Poincaré and the geometrization conjecture, says Carlson. "If not, it develops tools and ideas that will bring the geometrization conjecture into reach," he adds. "There is a tremendous amount of excitement."

A flying start

Perelman first made an impression in the United States more than a decade ago. The young Russian spent two years of his early career working at the Courant Institute, the State University of New York at Stony Brook and the University of California, Berkeley. "He was already considered extremely brilliant; this was apparent in conversation and on the basis of his work," says Jeff Cheeger, a mathematician at the Courant Institute. But Perelman had an unusual reputation, even then. "He had long hair and long fingernails, several inches long," remarks one colleague. "When someone asked him why, he said it was so he could open a book at the exact page he wanted."

Perelman's early work was impressive enough to garner several job offers from US universities. But he turned them down, returning to Russia and a research position with the Steklov Institute of Mathematics in St Petersburg. At that point, Perelman effectively disappeared — he stopped publishing papers or discussing his research with colleagues. "We would occasionally ask where he was," says a friend. "No one seemed to know what he was doing." Even people at the Steklov Institute didn't know what he was working on.

But there were hints that he hadn't gone off the mathematical rails. "I was in touch with him a little bit," says Cheeger. "Enough to see he was following some developments closely." But no one knew whether he was working on something brilliant, or if he had just burned himself out.

In 2002, Perelman revealed what he had been working on for eight years. In November, he posted a paper¹ on an Internet preprint server outlining a proof of the geometrization conjecture. He also sent e-mails to a few mathematicians, telling them that they might be interested in the manuscript.

"I took a look and found it very interesting. It was a very important paper, so I decided to

invite him here," says Gang Tian, a mathematician at MIT. At the same time, Perelman was invited to speak at Stony Brook. In the interim, Perelman posted a second paper³ detailing technical arguments for his proof.

Tangible benefits of this work are hard for non-mathematicians to grasp. So attention has focused on the Poincaré conjecture and speculation about the Clay prize — a million-dollar cash award that the Clay Institute offers to anyone who can solve one of the seven toughest problems in mathematics.

Maths over money

Perelman has refused to discuss the Clay prize with the media and doesn't talk about it with friends. "He never mentioned the prize," says Tian, who played host to the Russian when he visited MIT. "He was only interested in talking about mathematics."

To win the prize, Perelman's work must survive two years of scrutiny by his peers. But even then, it is not clear whether Perelman would accept the award. He refused an earlier prize granted by the European Congress of Mathematics for work he did in the early 1990s. His reasons for doing so were unclear.

What is clear is that Perelman does not care much for money. "We tried to take him to a nice restaurant in Boston," recalls Mrowka. "I think he'd rather have had his bortsch. In some senses he is refreshing because he's totally committed to mathematics."

Perelman shows similar ambivalence towards the many job offers he received after his US lecture tour. Although universities are reluctant to give details, they say that they never heard back about their offers. At a private dinner during his stay at MIT, several people tried to convince Perelman to work in the United States, but he insisted that he could not be tempted.

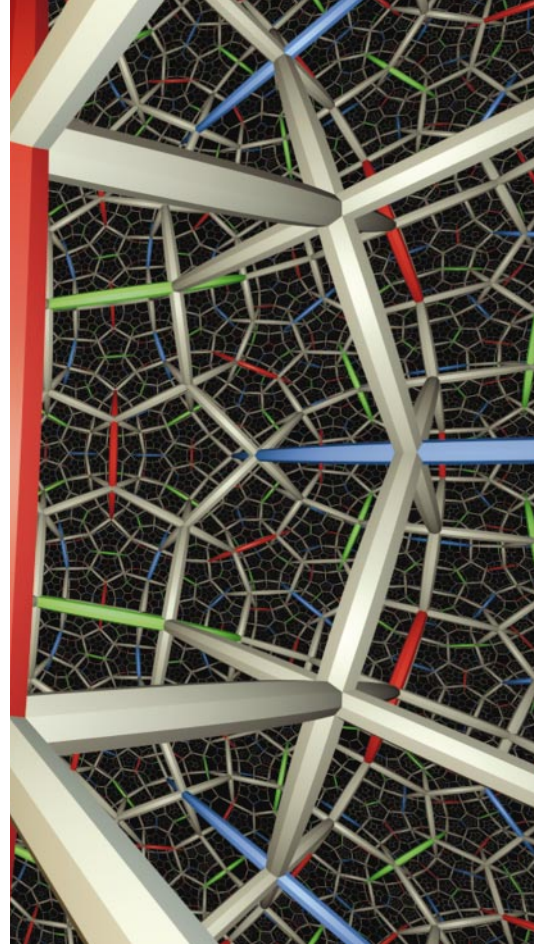
Tian admits that he doesn't understand why. "Maybe he doesn't want his peaceful life disturbed; maybe he is right, once he has proved he has a great theorem, that is much more valuable," he says.

For now, mathematicians seem content to study Perelman's work without him. Workshops have sprung up throughout the United States, and the Clay Institute has planned conferences on the topic. Tian has gone through a large part of the second paper and says that everything looks correct so far. He hopes to complete his examination by the summer.

The two years of community scrutiny will be up in 2005, and then we'll know whether Perelman's peers deem his work worthy of a Clay prize. But given his track record, Perelman is unlikely to pay the award much heed. ■

Emily Singer recently completed a short internship in Nature's Washington DC office.

1. Perelman, G. preprint at <<http://www.arxiv.org/abs/math.DG/0211159>> (2002).
2. Stewart, I. *Nature* **423**, 124–127 (2003).
3. Perelman, G. preprint at <<http://www.arxiv.org/abs/math.DG/0303109>> (2003).



C. CHANG/THE DAILY PRINCETONIAN



Grisha Perelman presents his proof, which covers Poincaré's conjecture, in a lecture last year.

itself. In 1904, the French mathematician Henri Poincaré argued that a three-sphere should follow the rule for a two-dimensional sphere — but he was unable to prove it.

The geometrization conjecture is a more general statement about three-dimensional 'surfaces', derived in the late 1970s by William Thurston, now at Cornell University in Ithaca, New York. According to Thurston, all three-dimensional surfaces are made from eight basic geometries, and he theorized that every shape can be described using these building blocks. Poincaré's unsolved conjecture is a limited case of Thurston's theory.

Richard Hamilton, a mathematician at Columbia University in New York, had previ-



The most important sexual organ

New evidence suggests that the brain begins to develop differently in males and females much earlier than was thought — before sex hormones come into play. Carina Dennis considers the implications.

Susan was whisked away as soon as she was born. Her mother wasn't allowed to bathe or undress her for several days; the doctors simply said that there was something wrong. Susan had an 'intersex' condition — her genitalia offered no clear indication whether she was a girl or a boy.

Although Susan is genetically male, doctors convinced her parents that she should undergo surgery to turn her into a girl — the operation is easier to perform, and medical opinion was that, with appropriate hormone treatment, she would develop normally in her assigned gender. And so it turned out: Susan is now in her twenties, married and happy as a woman.

But for a small proportion of the one in about 4,000 children born with 'ambiguous' genitalia, things don't work out so well. Tony was also born genetically male and had similar

surgery. But he never felt comfortable as a girl. After many years of unhappiness, he began testosterone treatment and now lives as a man.

Susan's and Tony's doctors believed that the brain is a sexual blank slate until it comes under the influence of sex hormones. But biologists are now starting to realize that hormones aren't the only significant determinant of the brain's sexual destiny. Indeed, male and female brains may even start moving down different developmental paths before sex hormones are produced in significant quantities. "There is plenty of evidence that hormones organize the brain sexually, but it's not the whole story," says Eric Vilain, a geneticist at the University of California, Los Angeles (UCLA).

New findings about the genetic and other factors that influence the brain's sexual development could do more than simply

rewrite the textbooks. They might provide insights into conditions such as transsexualism — and perhaps eventually lead to tests that could determine whether a baby with an intersex condition is more likely to grow up thinking, feeling and behaving like a man or a woman.

Sex discrimination

The dogma that hormones are the only important players in the brain's sexual development stems from early experiments showing that castrated male rabbit fetuses seem to develop as females, and newborn female guinea-pigs exposed to testosterone come to look and act like males. These results indicated that the 'default' sex in mammals is female, and suggested that male brain anatomy and behaviour develop as a consequence of exposure to testosterone.

More recently, geneticists uncovered the 'master switch' that kick-starts the development of the testes in mammals — a gene called *Sry* on the male-specific Y chromosome¹. But once this engine is running, testosterone produced by the testes was still assumed to direct the development of male characteristics. And so the idea that the differences between male and female brains are entirely a consequence of their exposure to sex hormones persisted. "The field was completely dominated by hormone people," says Vilain.

The realization that things are rather more complicated comes in part from studies of birds. Their sex chromosomes are different from those in mammals; and in the feathered world, the default pathway for sexual development is male rather than female. But again, the differences between male and female brains were thought to be simply a result of their exposure to sex hormones.

Convincing evidence that this isn't the complete story comes from a rare, naturally occurring zebra finch that was genetically male on the right side of its body, with bright plumage and a testis, and genetically female on the left, with dowdy feathers and an ovary. If brain sex depended solely on hormones, you'd expect both sides of the bird's brain to be the same, as they were both exposed to the same mix of male and female hormones coursing through the blood. But when researchers led by neurobiologist Arthur Arnold at UCLA examined the bird's brain, they found that the neural circuits that control male song were much larger on the right side².

All in the mind

Manfred Gahr, a neuroscientist at the Free University of Amsterdam in the Netherlands, created his own sexually confused birds. He took Japanese quail embryos and surgically switched their forebrains — which control adult sexual behaviour — before their gonads began to develop. If conventional wisdom is correct and sex hormones produced by the gonads do direct the brain's sexual development, then having a forebrain that was genetically the 'wrong' sex should have made no difference to the birds' adult sexual behaviour. And for the females given male forebrains, this was the case: they looked and acted like normal female quail. But the male birds

given female forebrains neither crowed to attract mates nor mounted females: their transplanted female brain tissue could not support typical male reproductive behaviour³. Their testes also failed to develop normally, indicating that — at least in male quail — a genetically male brain is required to complete the development of the gonads.

It's hard to do similar experiments in mammals, where embryological development occurs in the womb. But Vilain has used DNA microarrays to study gene activity in the brains of male and female mouse embryos from an early stage of development. Of the 12,000 genes active in the brain, 51

of Cancer Research in London. She now plans to compare mice that have been engineered so that *Sry* is active only in their testes with normal animals in which the gene is expressed in both brain and gonads, looking for any differences in behaviour and brain anatomy.

Under the influence

Other researchers are also working with mice in which *Sry* has been manipulated: male mice lacking *Sry* develop as females, whereas genetically female mice given *Sry* grow up as males⁸. Using such mice, Arnold teamed up with Ingrid Reisert at the University of Ulm in Germany to investigate a phenomenon that Reisert had identified more than a decade ago: the fact that some cells extracted from the midbrains of male and female embryonic rodents develop differently when grown in culture⁹. They found that these differences aren't simply a result of *Sry* kick-starting the production of testosterone in the testes — other genes on the sex chromosomes are also apparently involved¹⁰.

The animals used by Reisert and Arnold were created in the lab of Paul Burgoyne, a developmental biologist at the



Determined effort: does the fetal brain already know its sex before its gonads develop?

showed different levels of expression in the brains of male and female mouse embryos before the gonads had formed⁴. This suggests that, in mammals, male and female brains start travelling down different developmental paths from the outset, before hormones have entered the picture.

Whether the genes Vilain has identified have an important biological function in the brain remains unknown. To find out, he is now creating 'knockout' mice in which the genes have been systematically inactivated, to see what effect this has on brain development and sexual behaviour. Initially, Vilain intends to focus on genes on the Y chromosome. "These are obvious candidates," he says.

Other researchers are looking at *Sry* to determine whether it has a direct influence on the development of the brain, in addition to its role in the testes. The gene certainly seems to be expressed in the brains of male mammals⁵⁻⁷, but again, it's unclear what this means. "Expression patterns suggest where genes are acting, but we need functional studies to prove it," cautions Amanda Swain, a developmental biologist at the Chester Beatty Laboratories of the Institute

UK Medical Research Council's National Institute for Medical Research in London, who has engineered a range of mice with unusual configurations of sex chromosomes. "You name it, we can make it," says Burgoyne, who similarly aims to use the mice to tease out genetic influences on the way in which male and female brains are wired.

Studies of the non-hormonal influences on the sexual development of the mammalian brain are at an early stage. But Vilain is already trying to apply this emerging knowledge to the phenomenon of transsexualism — where individuals are physically and genetically one sex but perceive themselves to be members of the other. Reports of the condition sometimes running in families¹¹ have led some researchers to speculate that genetic factors may be involved. "Some of the genes differentially expressed in the brains of male and female mice may be altered in transsexuals," Vilain suggests.

He is now collaborating with Vincent Harley, a molecular biologist at Prince Henry's Institute of Medical Research at the Monash Medical Centre in Melbourne, which runs the largest transsexual clinic in Australasia. They are looking for variants in

the sequence of these genes to see if some are more common in transsexuals than in the general population. Their initial analysis will focus on three genes, located on the sex chromosomes, that are known to encode proteins that regulate the activity of other genes.

Such research into the genetics of gender identity is inevitably controversial — especially given the dark history of the eugenics movement, which viewed conditions such as transsexualism as ‘diseases’. But Harley has been overwhelmed by the positive response from the transsexual community since the study began. “I’ve had people e-mailing me to offer their help,” he says. “There is a stigma that transsexualism is a lifestyle choice. It may be very liberating to them if we can show that there is a genetic or biological basis.” Indeed, Harley has been invited to give talks about his research to transsexual groups in Melbourne.

Difficult decisions

Vilain also hopes that this research might result in tests that could help to predict whether a baby with an intersex condition is more likely to grow up feeling like a man or a woman. “Gender is not as simple as X and Y chromosomes,” he says. Many genes are likely to be involved and, if genetic variants associated with the gender that a person feels they belong to can be identified, then it may one day be possible to take a blood sample from a young intersex patient to determine which of these DNA sequences they carry. “If there was a biological tool that could predict the likely gender of a person based on their DNA sequence, that would be very useful for intersex children,” Vilain says.

Babies born with ambiguous genitalia are usually operated on within the first few months of life. “The majority of our patients are satisfied with their assigned gender,” says Garry Warne, a paediatric endocrinologist at the Royal Children’s Hospital in



Like its body, this zebra finch’s brain was ‘male’ on the right and ‘female’ on the left, hinting that more than sex hormones guided its development.

Melbourne, Australia. “In less than 8% of cases is the outcome bad.”

But when the wrong choice is made, as in Tony’s case, the consequences are severe. For as long as Tony can remember, he felt like a boy. In his thirties, after voicing his plight on Australian national television, he was finally able to get doctors to provide him with testosterone injections. Not surprisingly, Tony deeply resents the drastic surgery that cast him as a female.

Such distressing experiences lie at the

heart of an unresolved debate over when surgery should be used on intersex patients. Tony believes that it should be left until the individual can make an informed decision. But most clinical experts argue that early intervention avoids the psychological damage suffered by a child growing up with ambiguous genitalia. Warne, for instance, has surveyed intersex individuals in Vietnam and India — for whom surgery was not available — and found that most wished they had been operated on as infants.

Given the growing realization that hormones don’t act alone in directing the brain to develop along male or female lines, Vilain suggests that any surgery should be delayed until the child can offer consent and has started displaying gender-specific behaviour. But he hopes that it may one day become possible to use genetic tests to allow decisions about gender assignment to be taken with more confidence at an earlier age.

Tony, however, has mixed feelings about the use of genetic information to manage the treatment of intersex patients. “If it can lower the chances of the wrong outcome, that would be fantastic,” he says. “But that information shouldn’t be used to surgically reinforce one gender.”

Carina Dennis is Nature’s Australasian correspondent.

1. Koopman, P., Gubbay, J., Vivian, N., Goodfellow, P. & Lovell-Badge, R. *Nature* **351**, 117–121 (1991).
2. Agate, R. J. et al. *Proc. Natl Acad. Sci. USA* **100**, 873–878 (2003).
3. Gahr, M. *Proc. Natl Acad. Sci. USA* **100**, 7959–7964 (2003).
4. Dewing, P., Shi, T., Horvath, S. & Vilain, E. *Mol. Brain Res.* **118**, 82–90 (2003).
5. Lahr, G. et al. *Mol. Brain Res.* **33**, 179–182 (1995).
6. Mayer, A., Lahr, G., Swaab, D. F., Pilgrim, C. & Reisert, I. *Neurogenetics* **1**, 281–288 (1998).
7. Mayer, A., Mosler, G., Just, W., Pilgrim, C. & Reisert, I. *Neurogenetics* **3**, 25–30 (2000).
8. De Vries, G. J. et al. *J. Neurosci.* **22**, 9005–9014 (2002).
9. Beyer, C., Pilgrim, C. & Reisert, I. *J. Neurosci.* **11**, 1325–1333 (1991).
10. Carruth, L. L., Reisert, I. & Arnold, A. P. *Nature Neurosci.* **5**, 933–934 (2002).
11. Green, R. *Arch. Sex. Behav.* **29**, 499–507 (2000).



A question of gender: Vincent Harley (left) and Eric Vilain are hoping to find out whether there is a genetic basis for transsexualism.

Fraud offers big rewards for relatively little risk

We need to change the over-competitive culture that promotes publishing at all costs.

Sir — Your News Feature “It’s a scoop!” (*Nature* **426**, 222–223; 2003) thoughtfully explores the dangers of excessive competitiveness among researchers racing to be the first to publish high-profile scientific results. Although this can lead to secretiveness or sloppy work, a much worse possibility is that it is encouraging scientific fraud or other forms of misconduct.

For example, I asked a colleague who was familiar with the details of a recent case of scientific fraud why — given the risks — he thought the perpetrators had done it. He drolly observed that the real question is not why a few scientists commit fraud, but why more don’t do it. He went on to say that since the maximum penalty for getting caught (dismissal) was no worse than the routine penalty for not producing enough high-profile papers (no job), most junior scientists, at least,

have nothing to lose by committing fraud.

I suspect that, although blatant cases of fraud in high-profile journals are probably rare, more subtle forms may be quite common: skipping a control or two, claiming that the same result was obtained in a repeat experiment that was not made, or failing to mention that some essential cross-checks did not give the required result.

It is almost impossible to determine how often this kind of thing is happening, but that is not a cause for complacency. The notion that fraud will always be detected when the results are shown to be false assumes that the fraudulent results will always stand out sufficiently to attract scrutiny. Less blatant cases may pass unnoticed, or the researcher trying to push his or her work just a little might actually have guessed right. Furthermore, how many researchers these days have time to

repeat whole chunks of previously published work, or go back over older literature and critically examine every result for possible signs of fraud, let alone be willing to do anything about such suspicions?

What can be done? It probably requires changes in the way science is funded, and especially how junior scientists are employed. Some sort of international inspection body might help, similar to that used to find and punish cheats in the sporting world. And as individuals, we can all help to develop a scientific culture that discourages fraud or related misconduct as much as possible, while finding ways to increase the rewards for honest but effective competitors.

T. M. Fenning

Max Planck Institute for Chemical Ecology,
Beutenberg Campus, Hans-Knöll-Strasse 8,
D-07745 Jena, Germany

Biodefence funds have tight strings attached

Sir — From reading your News story on the trial of Thomas Butler, “Plague trial verdict leaves biologists split on defence” (*Nature* **426**, 593; 2003), it is not yet clear that US scientists are eager to work on biodefence projects, as suggested by the former president of the American Society for Microbiology, Ronald Atlas. What is clear, however, is that they are eager to get their hands on the flood of new federal money for biodefence research. No doubt the money will be used to finish current research projects, buy new equipment, hire technicians, fund graduate students and impress tenure committees. Only later will they have to consider the burden that comes with that money: regulatory paperwork, conflicting rules and regulations, and the scrutiny of the Federal Bureau of Investigation and other security bureaucrats.

Atlas says he has not seen evidence that new regulations are having an adverse impact on researchers, but that is largely because the ink is not yet dry on many of the new laws and regulations. We will have to wait a few years to see their full impact on academic research.

“We live in a new regulatory environment; it is our responsibility as scientists and citizens to comply with the laws and regulations,” writes Atlas. It is also our responsibility as scientists and citizens to question the utility and enforcement of such laws and regulations. As the

economist Lester Thurow reminds us (*Atlantic Monthly* **283**, 6; 1999): “being skeptical and refusing to accept authority are the secrets of scientific advancement.”

Edward McSweeney

1692 Barrister Court, Crofton,
Maryland 21114, USA

Bridging a know-do gap

Sir — Tikki Pang, in Correspondence (*Nature* **426**, 383; 2003) about your Editorial “In praise of Gates” (*Nature* **425**, 435; 2003), calls for the Gates Foundation to support the translation of knowledge into actions to improve people’s health.

Pang will be pleased to hear that the Gates Foundation is backing one key solution to addressing the ‘know-do’ gap: education. Among other projects, it is supporting an innovative ‘E-learning certification programme in global health’ directed at health professionals working in Africa, co-authored by African and international medical scientists and clinicians (see www.tall.ox.ac.uk/globalhealthprogramme). The aim is to provide research-based and up-to-date information where and when it is needed. The first module will be on malaria.

Sarah Davies*, Abiola Sulaimon†

*Technology-assisted Lifelong Learning,
Department for Continuing Education, University
of Oxford, Littlegate House, Oxford OX1 1PT, UK

†Nuffield Department of Clinical Medicine,
University of Oxford, John Radcliffe Hospital,
Oxford OX3 9DU, UK

Hibben was not proved guilty of misconduct

Sir — On behalf of the University of New Mexico (UNM), there are three points I wish to make in response to the News story “University buildings named on shaky ground” (*Nature* **426**, 374; 2003) about the UNM Center for Archeological Research and its benefactor Frank Hibben.

First, scientific controversy is the norm, rather than the exception, when it comes to research. It is quite different from scientific misconduct or fraud. Second, as noted in the News story, no allegation of scientific misconduct has ever been filed concerning Hibben’s work in any of the venues in which such charges can be filed. Third, in view of the previous points, resurrecting decades-old controversies eighteen months after Hibben’s death and characterizing them as fraud is unworthy of *Nature*. As your News story pointed out, “he was a mentor to many of today’s leading US archaeologists, and is widely credited with popularizing modern archaeology”.

We feel that these points and his many research and teaching accomplishments put our decision to name the Hibben Center in his honour on solid, not shaky, ground.

Terry L. Yates

Vice Provost for Research, University of New
Mexico, Albuquerque, New Mexico 87131, USA

Correction P. Kaaret’s Correspondence letter (*Nature* **427**, 287; 2004) should have included the expression ($\chi < 10^{-4}$), not ($\chi < 10^4$). *Nature* apologizes for this error.

Crops behaving badly

Are transgenic crops the reckless delinquents their critics claim?

Dangerous Liaisons: When Cultivated Plants Mate with their Wild Relatives

by Norman C. Ellstrand

Johns Hopkins University Press: 2003.

268 pp. \$65, £48

Rick Roush

The ecological effects of genetically modified (GM) crops remain controversial, despite evidence of the crops' agricultural benefits, such as reduced pesticide use, fewer human poisonings and increased net incomes for farmers. Their most ardent critics argue that GM crops lead to the rapid evolution of resistance in pests, harm non-target species and soil organisms, and create 'superweeds' by introducing transgenes into wild plant populations.

However, the transgenic crops that have been commercialized to date seem largely to have defied these predictions. For example, there has been no detectable increase in insects' resistance to the bacterial toxin produced as a pesticide by transgenic *Bt* crops. Resistance to the herbicides used with GM crops (mostly glyphosate) is still relatively minor compared to that with other widely used herbicides. The damaging effect of *Bt* crops on Monarch butterflies proved to be exaggerated, and studies on other non-target species have consistently failed to find any significant direct effects. Studies of the interactions of soil organisms and processes with GM crops have also failed to show any significant detrimental effects.

In *Dangerous Liaisons*, Norman Ellstrand addresses the risk of GM crops sharing their genes with wild and weedy plants — either uncultivated populations of the crop itself or closely related species. His well-written book is aimed at a wide audience of students, academics and policy-makers. The popular press seems to have a difficult time with stories about pollen flow from GM crops, typically extrapolating from studies on sunflowers, maize, canola or sugar beets to all GM crops grown anywhere. Ellstrand's book will help them identify more specific risks. The book will surely gain wide attention because of its sexy title and emphasis on GM crops, but I was more impressed by Ellstrand's documentation of the broader effects of conventional agriculture in swamping wild populations of rare plant species.

Ellstrand provides an introductory section for readers who are not population geneticists, before detailing hybridization between domesticated plants and their wild relatives, and then presenting his interpretation of these observations.



Playing the field? Hybridization between transgenic and wild maize is a potential problem in Mexico.

Despite his effort to provide this broad context for hybridization between crops and wild plants, and its consequences, I suspect that most readers will focus on chapter 7, where Ellstrand contrasts his views with an opening quote from the Israeli plant scientist Jonny Gressel: "Most crops have no interbreeding relatives in most of the world." Ellstrand reviewed the data for the world's 25 most widely planted crops, summarized their interbreeding with wild plants in a single table, and showed that 22 of them do hybridize with wild relatives somewhere in the world. I suspect that this table will be the most widely referenced in the book, and wish that a few of the distributions were more precisely stated. For example, cotton, beans and potatoes are listed with a "multi-continental" distribution of hybridization; more precisely this refers to Latin America (and, for cotton, some islands in the Caribbean and Pacific).

But what about Gressel's proposition, especially in the context of GM crops? Using statistics from the database of the Food and Agriculture Organization of the United Nations, cited by Ellstrand (<http://apps.fao.org>), I checked on the dominant GM crops. At least 87% of the world's soybean crop, and 95% of the world's maize and cotton, are grown in countries for which Ellstrand lists no hybridization — and even for those countries with hybridization, such as China for soybeans, wild relatives are found in only some areas. Many of the other crops are complicated to tabulate, but Gressel

seems to be correct that most crops are not interbreeding locally with wild relatives.

This still leaves the possibility that serious problems could arise in the few areas of the world where hybridization can occur. This currently seems possible for GM canola in Canada and the United States, and for transgenic maize that is probably growing illegally in Mexico (but has apparently escaped documentation in the refereed literature). But even after reading this book, I haven't seen any evidence of harm to human health or to the environment (including weediness) from such hybridization. Where are the superweeds that were predicted to occur from the exchange of transgenes with wild relatives?

In contrast to the lack of evidence for deleterious effects of gene flow from GM crops, there is evidence that conventional agriculture has adversely affected wild plants through genetic swamping of their populations, and that wild plants have generated weediness in crop-weed hybrids. As noted by Ellstrand, "problems associated with hybridization between conventional crops and their wild relatives received scant attention until potential gene-flow problems were described for transgenic crops". For example, hybridization with cultivated rice has been implicated in the near-extinction of an endemic Taiwanese wild rice. Hybridization of maize with its ancestor teosinte may be contributing to the extinction of teosinte populations. Indigenous cotton in the Galapagos Islands could be at risk of extinction or replacement as a result of hybridization

with cultivated cotton. Ellstrand cites similar evidence for at least another nine species. He also documents in great detail the history of sugar beets in Europe, where hybrids between cultivated beets and their progenitors, the sea beets, have caused major weed problems.

Everyone interested in the effects of cropping on plant biodiversity, the evolution of weeds and the risks of GM crops should read this book. Critics and supporters of transgenic crops will continue to debate whether the relatively benign environmental and agronomic disadvantages of GM crops have been due to largely to luck or to an adequate regulatory system. Ellstrand reminds us in detail that the reliability of future successes depends on more careful risk assessment of hybridization with wild relatives. ■

Rick Roush is director of the University of California Statewide Integrated Pest Management Program, based at University of California, Davis, California 95616, USA.

Japan's secret weapons

A Plague Upon Humanity: The Secret Genocide of Axis Japan's Germ Warfare Operation

by Daniel Barenblatt

HarperCollins: 2004. 256 pp. US\$25.95, Can\$35.95. To be published in the UK in March 2004 by Souvenir Press, £20

Alastair Hay

When Japan invaded Manchuria in 1931, it alleged Chinese involvement in blowing up a section of the Japanese-owned South Manchurian Railway. The charge was false: confessions in 1945 by some of those involved confirmed that Japanese engineers carried out the attack to provide the pretext for the invasion of the Asian mainland. History is littered with similar pretexts.

Condemnation of Japan's invasion by the League of Nations followed relatively swiftly, prompting Japan to leave the body in 1933. Japan subsequently attacked China in 1937 and the US naval base at Pearl Harbor in 1941, and then invaded Vietnam, Malaysia and the Philippines. The brutality of the invaders has been well documented and the success of the military campaign waged by Japan is now textbook material.

Less well known are some of the means that Japan used to get its way, and biological warfare is just one of these. Daniel Barenblatt's *A Plague Upon Humanity* is an attempt to redress this ignorance and to provide a readable book that documents what happened. Barenblatt is bemused, asking himself how it is that the "startling" information about Japan's use of biological warfare "is

not common knowledge, as is the Holocaust and the experiments of the Nazi doctors?"

A Plague Upon Humanity addresses this question, the succinct answer to which is political expediency. Following the defeat of Japan in the Second World War, Japanese scientists and doctors involved in the biological-warfare programme did a deal with the United States. The arrangement was that the United States would receive details of the programme and results of experiments, and in return the Japanese researchers would receive immunity from prosecution. The deal only became public knowledge in 1980.

As far as the United States was concerned, the information it received was a windfall. Not only had the Japanese used plague, cholera, paratyphoid and anthrax in attacks on Chinese civilians, but they had also carried out experiments on some 3,000 people. These experiments are some of the most gruesome recorded, and many involved deliberately infecting people. Individuals — primarily locals who had infringed the Japanese penal codes — were exposed to a single bacterial or viral agent and monitored. As the disease progressed, those infected (the *maruta*, or 'logs of wood', as they were often called) were assessed on a daily basis; official records show that some were even operated on while still alive. But none survived their ordeal — all 3,000 either succumbed to the infection or were killed and their tissue retained.

The Japanese tested at least 19 bacteria and viruses as candidates for biological warfare. Writing about this in 1947, Edwin Hill, chief of basic sciences at Camp Detrick in the United States, the main research centre for biological warfare, noted that his government had obtained the data for a "mere pittance by comparison with the actual cost of the studies". The value of the data was not in question for, as Hill observed: "Such information could not be obtained in our own laboratories because of scruples attached to human experimentation." Hill does not seem to comment on the scruples of those who use the information or who failed to prosecute the scientists for war crimes.

None of the above is new information, and most of it is retold in Barenblatt's book. Drawing heavily on already published works, particularly *Factories of Death* by the late Sheldon Harris and *Unit 731* by Peter Williams and David Wallace, this latest addition to the literature is something of a disappointment. Barenblatt is right to be outraged about his subject matter and the fact that it still remains relatively obscure; perhaps his book will reach a wide audience and fulfil one of his aims. But for researchers in this field, his book will be a let-down. The sources of the many facts cited in the text are not referenced. Given that the material is so outrageous, it is all the more important to be scrupulous about documenting its origins.

Japan's biological-warfare programme was the brainchild of Shiro Ishii, an immunologist who inveigled his way into Japan's military hierarchy and pushed at an open door to get the funding he needed. His supporters were persuaded that biological warfare would assist Japan's imperial ambitions, enabling Ishii to set up shop in Manchuria at a centre known as Unit 731. Barenblatt demonizes Ishii, and there is much to condemn, but the actions of this rogue scientist are not new. What is missing from Barenblatt's book, and what is sorely needed, is solid evidence of the number of Chinese who died as a result of Japan's use of biological warfare in the field. Sadly, there are few Chinese scholars working in this area with access to the necessary archives.

Barenblatt refers to Chinese sources who claim that as many as 580,000 died as a result of Japan's biological warfare. The figure is said to be preliminary: the number may increase as house-to-house enquiries by investigators turn up more victims. Whether these deaths are truly attributable to the biological-warfare programme, or are the result of insanitary conditions brought about by mass population movements in war, is not clear. What is beyond dispute, however, is that Japan used biological warfare to kill many people. The more people who know this, and about those who did deals to suppress the information, the better. ■

Alastair Hay is professor of environmental toxicology, School of Medicine, University of Leeds, Leeds LS2 9JT, UK.

Love, actually

Why We Love: The Nature and Chemistry of Romantic Love

by Helen Fisher

Henry Holt: 2004. \$25, 288 pp.

Alison Jolly

Tristan is on the road with Iseult, who is planning to marry his uncle, Mark. They shoot up a potion of dopamine and nor-adrenaline enhancer, take one look at each other and fall in love. A serotonin suppressor ensures that the passion becomes obsessive. Delirious with joy, Tristan answers an advertisement from his university at Rutgers, New Jersey, recruiting lovestruck undergraduates to have their brains scanned. In the analysis, his caudate nucleus shines out, a part of the ancient reptilian brain tuned to anticipate and discriminate between rewards. So too does his ventral tegmental area, producing ever more dopamine to drench his emotions. As Helen Fisher writes: "When I first looked at those brain scans, with the active brain regions lit up in bright yellow and deep orange, I felt the way I feel on a summer night when



The chemistry's right: it was love at first sight for Richard Burton when he met Elizabeth Taylor.

I gaze at the sparkling universe: overwhelming awe."

Fisher's populist book attempts to explain the mechanism of romance. It is snappily written and filled with quotations. Here's Richard Burton's first sight of a 19-year-old Elizabeth Taylor: "She was so extraordinarily beautiful I nearly laughed out loud. She... was famine, fire, destruction and plague... Her breasts were apocalyptic, they would topple empires before they withered... those huge violet eyes... had an odd glint... Aeons passed, civilizations came and went while these cosmic headlights examined my flawed personality. Every pockmark on my face became a crater of the moon."

Fisher's thesis is that headlong infatuation is a universal human trait. The brain scans she has made of volunteers in and out of love, and their chemical background, are the scientific kernel of the story. The book's aim is to interest people who are fascinated by love found or love lost, with a light hand on the science side.

What interests me is what Fisher does not say. She makes a deliberate choice to ignore the mind-brain and evolution-culture controversies of recent decades. She does not bother justifying her approach. She quotes sociobiologists and evolutionary psychologists as sources, with no mention that they still raise many people's hackles. She simply describes a mating system that sets our emotions on fire at the sight and touch of a beloved individual, but where lust, companionship and romance do not map perfectly onto each other. Then she speculates how this system enables the successful propagation of a species that needs commitment

between partners, at least during the first four years of a dependent child's life. So far, I am with her. It is high time we outgrew the phlogiston theory of biology-free humans.

However, she also chooses to ignore the ethics of her studies. This worries me more. She accepts a chemical-based society. When love is lost, depression sets in. In mild form, it is adaptive, enabling the jilted lover to let go. In extreme forms, it is horror. Fisher reports that currently "some 7.1 million Americans take serotonin boosters to counter depression, stress, bereavement, or the despair of tragic love". She doesn't strongly advise that you join them, but points out that anything is better than suicide. When you are ready to fall in love again, she suggests reviewing your medication. Serotonin boosters could block romantic obsession, trivializing your next relationship.

I conclude that King Mark does not stalk Iseult or menace Tristan, her knightly lover. His first love potion having gone astray, he simply doses Iseult with an antidote before readministering the potion, making sure this time that she looks at him, not Tristan. Tristan is still obsessed by old-fashioned romantic love, ignoring its chemical bases. He sneaks into the castle disguised as a jester. His faithful hound recognizes his smell, but not only does Iseult not recognize him, she does not even care. Now you can decide the story's end. Does Tristan die of a broken heart? Or does King Mark don a doctor's white coat and advance on his rival, syringe in hand — not to murder Tristan, but to erase his love? ■
Alison Jolly is in the Department of Biology and Environmental Sciences, Sussex University, Brighton BN1 9QG, UK.

Fossils off the record

To See the Fellows Fight: Eye Witness Accounts of Meetings of the Geological Society of London and its Club, 1822–1868

edited by John C. Thackray
British Society for the History of Science
(Monograph no. 12). 2003. 244 pp.
£15, \$26

Martin Rudwick

The Geological Society of London, founded in 1807, was one of the first learned societies to be devoted to a specific science, and it was the first in the world to be devoted to geology. The most significant feature of its early meetings was one that nowadays we take for granted: it was agreed to allow discussion of the papers that were read.

This was in stark contrast to what was customary at the Royal Society, and at learned societies generally, where comments were relegated to private conversations after the formal meeting had ended. There was a fear that if discussion were allowed within a meeting it might degenerate into abusive argument, and that this would undermine the public image of the sciences as bodies of factual knowledge, about which there could not properly be divergent opinions.

In the event, however, the Geological Society's decision to allow discussion was vindicated, and other scientific bodies eventually adopted the same custom. It signalled a tacit acceptance of the vital role of argument and disagreement in the process of establishing and improving scientific knowledge. The Geological Society's meetings became famous for their lively arguments. As the editor of the Tory intellectual *Quarterly Review* commented: "I don't much care for geology, but I do like to see the Fellows fight."

This memorable quote has been used as the title for the late John Thackray's valuable collection of contemporary reports of the meetings of the Geological Society in its early decades. The brief official record of the titles of the papers read is followed by excerpts from letters and private journals reporting or commenting on the discussions or on the papers themselves. No formal account of the discussions was published, so these are the only evidence of what those present thought of the papers. Some other valuable accounts are of discussions at the society's informal dining club, rather than its formal meetings.

The collection starts in 1822, with an occasion that epitomizes the star-studded period in the history of English geology that had by then begun. William Buckland of Oxford University described at the club's dinner table how he had reconstructed the ecosystem of 'ante-diluvial' (Pleistocene)



Wild speculation? William Buckland won a Copley Medal for his work on the Pleistocene.

Yorkshire, based on fossils recently found in Kirkdale Cave: "The hyaenas, gentlemen, preferred the flesh of elephants, rhinoceros, deer, cows, horses &c., but sometimes unable to procure these & half starved they used to come out of the narrow entrance of their cave in the evening down to the water's edge of a lake which *must* once have been there, & so helped themselves to some of the innumerable water-rats in which the lake abounded." The prosaic young Charles Lyell was baffled: "We could none of us discern how far he believed himself what he said," he reported to his friend Gideon Mantell, who had not been present. (In fact, Buckland had good evidence for his reconstruction, and the Royal Society gave him a Copley Medal for it.)

The last report comes from 1868, when summaries of the discussions at meetings began to be published in the society's *Proceedings*. By this time, the brilliance of the earlier period had begun to fade.

Thackray's book should not be read from cover to cover, but random dipping will yield some absorbing and often entertaining vignettes of a lively and argumentative bunch of highly intelligent scientists. And for historians the book is an invaluable aid to further research, as it brings together comments that illustrate the character and development of some important controversies, including those about the Devonian system and the Pleistocene glaciations. We should be grateful to those of John Thackray's friends who have worked hard to bring his anthology to publication; it is a fine memorial to a colleague who was much loved and respected, and is greatly missed.

Martin Rudwick is in the Department of History and Philosophy of Science, University of Cambridge, Cambridge CB2 3RH, UK. His most recent book is Georges Cuvier, Fossil Bones and Geological Catastrophes (1997).

Science in culture

Animated animals

Jean-Baptiste Oudry's paintings brought beasts to life.

Martin Kemp

How should we portray an animal to provide information to those who are unfamiliar with it? It depends on what sort of information is to be conveyed. An illustration concerned with identification will pay attention to shape and detail, treating the subject like a static specimen. Animal behaviourists will require that action be given priority — even expression, as in Darwin's *The Expression of the Emotions in Man and Animals* (1872). Ecologists will look to the setting of the animals in their environments. But answers to this question have also changed over the centuries.

Before the eighteenth century, picture books of the world's fauna generally displayed animals as specimens, rarely in action. Painters, by contrast, typically set appropriate animals in stirring narratives, such as lion hunts, or within other pictures, such as portraits of aristocrats. The greatest painter of animals in the eighteenth century, Jean-Baptiste Oudry (1686–1755), not only portrayed animals with exceptional vitality in traditional genres, but also changed the way that 'zoological' specimens could be characterized.

In the 1730s, Oudry was commissioned by Louis XV's surgeon, Francois Gigot de La Peyronie, to undertake a series of paintings, many life-sized, of the prized exotics in the Royal Menagerie in Versailles, France. The paintings were apparently intended for the royal collection and were to be engraved. But when La Peyronie died in 1747, the project lapsed and the pictures remained in the artist's hands. They were scooped up by the Duke of Mecklenburg-Schwerin, another of Oudry's patrons. Now, much of the series has returned to France from the Staatliches Museum in Schwerin, Germany, for a compelling exhibition at Versailles.

Based on wonderfully spirited drawings, the paintings of animals exude a sense of their living presence, their deportment, character and expression. They are action portraits, sometimes heroic, sometimes sentimental, sometimes amusing. There is, not surprisingly given that Oudry illustrated the *Fables* of Jean de La Fontaine, a strong anthropomorphic element to the paintings. But the anthropomorphism has a solid base, not unrelated to Darwin's seminal book, *The Descent of Man*.

Oudry described the *Male Leopard* (right), a painting exhibited at the Paris Salon in 1741, as "a male tiger [sic] in choleric

disposition". Its companion piece portrays "a female tiger in a tranquil attitude". These brief descriptions serve to key Oudry's portraits into the great tradition of characterizing the humours and temperaments in painting exemplified by the teaching of Charles Le Brun, founder of the French Academy. Animals, like people, were seen to exhibit generic characteristics: lions, like warrior-kings, were bold but just; foxes were cunning; mice were timorous; females maternal and docile. They also manifested the specific emotions of the moment.

Le Brun provided a series of physiognomic formulae with which to convey character and expression. His treatise on *General and Particular Expression* showed how such sentiments and feelings as love, jealousy, pain, anger, terror and sadness were displayed by the configuration of all of the features of the head. Oudry's *Male Leopard* is the feline embodiment of Le Brun's *Choler Mixed with Rage*. The eyebrows are arched, the brow is bunched, the eyes open wide, seeming to protrude from their sockets; the lips are drawn back to reveal the teeth; and nostrils and whiskers flare. The leopard's tail thrashes angrily, and it twists with lithe agility to confront an unseen adversary. Oudry's fizzing brushwork matches the vivacity of his draughtsmanship.

Although Darwin was sharply critical of Le Brun's formulaic tendencies, he did insist on the underlying universality of expressive mechanisms, many of which have survived in humans divorced from their functional contexts. Oudry's common factor was the doctrine of the humours; Darwin's was the descent of man from animals. Like all great artists, Oudry achieves an insight that transcends the doctrines of his times.

Martin Kemp is professor of the history of art at the University of Oxford and co-director of Wallace Kemp/Artakt.



Jean-Baptiste Oudry's *Male Leopard* reveals the beast's emotions.

Engineering complex systems

The emergent properties of complex systems are far removed from the traditional preoccupation of engineers with design and purpose.

J. M. Ottino

Complex systems can be identified by what they do (display organization without a central organizing authority — emergence), and also by how they may or may not be analysed (as decomposing the system and analysing sub-parts do not necessarily give a clue as to the behaviour of the whole). Systems that fall within the scope of complex systems include metabolic pathways, ecosystems, the web, the US power grid and the propagation of HIV infections.

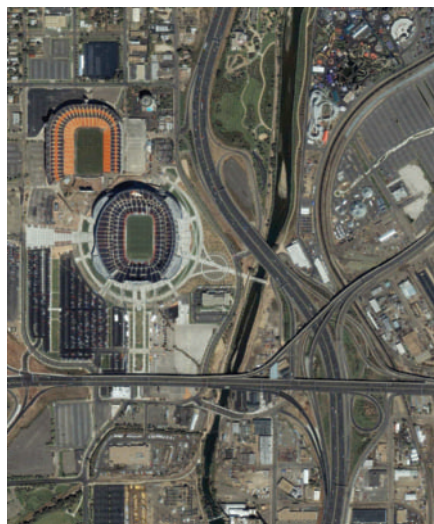
Complex systems have captured the attention of physicists, biologists, ecologists, economists and social scientists. Ideas about complex systems are making inroads in anthropology, political science and finance. Many examples of complex networks that have greatly impacted our lives — such as highways, electrification and the Internet — derive from engineering. But although engineers may have developed the components, they did not plan their connection.

The hallmarks of complex systems are adaptation, self-organization and emergence — no one designed the web or the metabolic processes within a cell. And this is where the conceptual conflict with engineering arises. Engineering is not about letting systems be. Engineering is about making things happen, about convergence, optimum design and consistency of operation. Engineering is about assembling pieces that work in specific ways — that is, designing complicated systems.

It should be stressed that 'complex' is different from 'complicated'. The most elaborate mechanical watches are appropriately called *très compliqué*, for example the Star Caliber Patek Philippe has over 1,000 parts. The pieces in complicated systems can be well understood in isolation, and the whole can be reassembled from its parts. The components work in unison to accomplish a function. One key defect can bring the entire system to a halt; complicated systems do not adapt. Redundancy needs to be built in when system failure is not an option.

How can engineers, who have developed many of the most important complex systems, stay connected with their subsequent development? Complexity and engineering seem at odds — complex systems are about adaptation, whereas engineering is about purpose. However, it is robustness and failure where both camps merge.

Consider the recent debate of the balance between performance and risk. Many systems



More than the sum of its parts: complex systems, such as highways, are constantly evolving.

self-organize to operate in a state of optimum performance, in the face of effects that may potentially destroy it. However, the optimal state is a high-risk state — good returns at the price of possible ruin. Most engineers are risk averse, and would prefer to eliminate the probability of catastrophic events. Recent work borrows concepts from economic theories (risk aversion, subjective benefit of outcomes) and argues that one can completely remove the likelihood of total ruin with minor loss of performance. This falls squarely in the realm of engineering, but the discussion has been driven by physics.

Engineers might also learn from social scientists. In social sciences, there is no such luxury as starting *de novo* — systems are already formed, one has to interpret and explain. Many engineering systems, such as the web or the US power grid, also fall into this category. How will they behave? How robust are they? How might they fail?

Although systems where self-organization has already happened present challenges, there are also opportunities in situations where self-organization can be part of the design. Could we intelligently guide systems that want to design themselves? Is it possible to actually design systems that design themselves in an intelligent manner? Self-organization and emergence have been part of materials science and engineering for quite some time, after all, lasers and superconductivity depend on collective phenomena. Emergent properties should strike a chord in materials processing, and also in the nanoworld. At larger scales, there is already

work in directed self-assembly and complex dissipative systems, which organize when there is energy input. However, practical processing by self-assembly is still not a reality, and there is work here for engineers.

But the choice need not be just between designing everything at the outset and letting systems design themselves. Most design processes are far from linear, with multiple decision points and ideas 'evolving' before the final design 'emerges'. However, once finished, the design itself does not adapt. Here, engineers are beginning to get insight from biology. The emergence of function — the ability of a system to perform a task — can be guided by its environment, without imposing a rigid blueprint. For example, just like the beaks of Darwin's finches, a finite-element-analysis of a component shape such as an airfoil can evolve plastically through a continuum of possibilities under a set of constraints, so as to optimize the shape for a given function.

Engineers calculate, and calculation requires a theory, or at least an organized framework. Could there be laws governing complex systems? If by 'laws' one means something from which consequences can be derived — as in physics — then the answer may be no. But how about a notch below, such as discovering relationships with caveats, as in the ideal gas 'law', or uncovering power-law relationships? Then the answer is clearly yes.

Advances will require the right kinds of tools coupled with the right kind of intuition. However, the current engineering courses do not teach about self-organization, and few cover computer modelling experiments.

Despite significant recent advances in our understanding of complex systems, the field is still in flux, and there is still a lack of consensus as to where the centre is — for some, it is exclusively cellular automata; for others it is networks. However, the landscape is bubbling with activity, and now is the time to get involved. Engineering should be at the centre of these developments, and contribute to the development of new theory and tools. ■

J. M. Ottino is at the R. R. McCormick School of Engineering and Applied Sciences, Northwestern University, Evanston, Illinois 60208, USA.

FURTHER READING

Ball, P. *Critical Mass* (Heinemann, Portsmouth, 2004).

Barabási, A.-L. *Linked: The New Science of Networks* (Perseus Publishing, Cambridge, 2002).

Hartwell, L. H. *et al. Nature* **402**, (suppl.) C47–C52 (1999). Center for Connected Learning and Computer-Based Modeling

► <http://ccl.northwestern.edu/netlogo>

Fungi and the food of the gods

Keith Clay

Plants protect themselves against attacks by microorganisms in various ways. In some circumstances at least, it turns out that they enlist the help of mutualistic fungi in defence of the home front.

Chocolaholics will be pleased with the results of Arnold *et al.*¹, published last month in *Proceedings of the National Academy of Sciences*. In the long term, the findings promise a strategy for increasing production of the crop that gives rise to this delectable treat. More immediately, however, they will come as a revelation to many plant biologists. They show that a poorly understood group of microbes provide their host plants with an effective defence against pathogens. The microbes are 'endophytes', fungi that live inside leaves without causing disease, and they appear to be ubiquitous associates of all plants.

Cacao provides the raw material for chocolate, the botanical name for the source tree (Fig. 1) being *Theobroma cacao* (which literally means 'food of the gods'). Production occurs primarily in the tropics of the Old World, owing to the high disease prevalence in the cacao tree's native American tropics. Arnold *et al.* report on endophytes and disease incidence in cacao trees in Panama, and find that the trees support diverse communities of fungi that are distinct from the endophytes inhabiting other tree species. The team also experimentally manipulated fungal endophyte colonization in cacao seedlings before infecting them with a virulent pathogen. They found that endophyte inoculation significantly reduced the incidence of disease and its damaging effects.

It is well known that a great variety of microorganisms associate with plant roots, so it should not be surprising that similar diversity exists in the above-ground parts of plants. Endophytic fungi have been isolated from virtually all plant samples examined^{2,3}. But what, if anything, are these microbes doing? The best-understood case to date has been that of endophytes of tall fescue, *Festuca arundinacea*, and other cool-season grasses⁴ (so-called because they tend to grow best in a cool, moist climate and are usually winter hardy). These fungi are vertically transmitted — that is, from generation to generation through seed. They infect the plant as a whole, and the host grass gains some protection from grazing animals as a



Figure 1 The cacao tree, *Theobroma cacao*. The pods evident in this photo contain the beans from which chocolate is made. Inset, experimental cacao plants showing the effects of disease on the leaves of plants that have been inoculated with endophytes (centre) and the much greater damage to leaves of control plants.

result of alkaloid toxins that the endophytes produce.

This example is not generally representative, however. In most plants, it is individual leaves that are colonized by a diversity of endophytic fungi, each forming highly localized infections. These infections are acquired independently by horizontal transmission through wind- or water-dispersed spores. In grasses, the persistence of highly mutualistic fungi is evidently a direct consequence of their strict vertical transmission through seeds. The diversity and dynamics of horizontally transmitted endophytes, or other mutualistic microbes, is less clear⁵. Given that endophytes use their hosts' resources, they must entail some cost to the plant. If the costs outweigh the benefits, why don't plants defend themselves against infection? And if the benefits outweigh the costs, what are these fungi doing to help the plant?

Arnold and colleagues¹ examined the

endophyte community of cacao by isolating the fungi from surface-sterilized leaf fragments on agar plates. They found that all samples were highly infected, with infection approaching 100% as the leaves aged (Fig. 2, overleaf). They isolated an average of 10 distinct taxonomic groups of endophyte from each leaf and a total of 344 groups from 126 leaves. The number of different endophytes per leaf also increased with leaf age. These findings agree with earlier results of Arnold *et al.*³ showing the high diversity of endophytic fungi in tropical trees. Together, their observations indicate a staggering level of fungal biodiversity (see also ref. 6), although other studies^{7,8} suggest more modest levels of diversity. If the cacao seedlings were protected from aerial deposition of spores, endophyte colonization fell to less than 1%, demonstrating that the fungal community develops over time by horizontal transmission. There was no evidence of the vertical

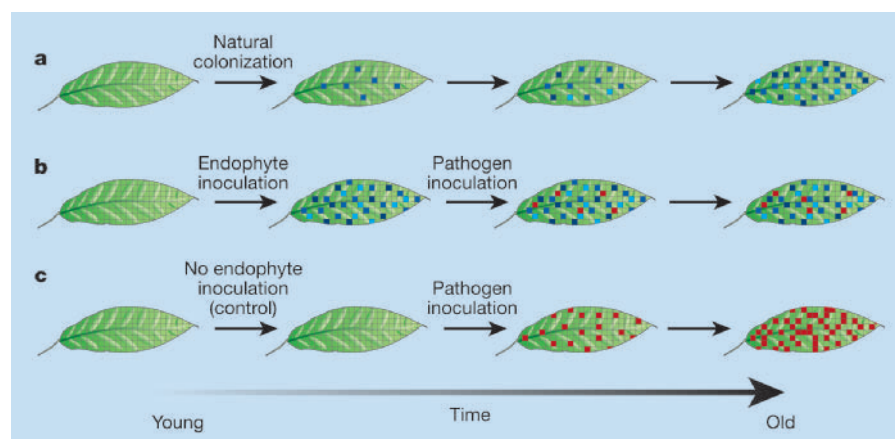


Figure 2 The colonization of leaves by endophytes, and pathogen deterrence. **a**, The natural colonization of leaves, increasing over time, by a diverse endophyte community. **b**, **c**, The results of the key experiments carried out by Arnold *et al.*¹ on sterile cacao seedlings. **b**, Pathogen attack of plants that had been inoculated with endophytes resulted in comparatively little pathogen occupation of leaf cells and limited damage. **c**, By contrast, pathogens became well established in endophyte-free control leaves and caused severe damage or leaf death.

transmission through seeds that is seen in grasses.

Arnold *et al.*¹ next looked at the distribution of endophyte diversity in cacao. The most similar communities occurred in neighbouring sites, and they became more divergent as the distance between sites increased. There was little evidence that endophyte composition reflects habitat type. The authors also sampled three tree species, including cacao, at a single site and found that most types of endophyte showed host affinity, infecting only one of the three. In fact, endophyte communities from different host species at a single site were more divergent than communities from cacao trees at separate sites. Arnold *et al.* confirmed this apparent host preference using agar plate assays in which fungal endophytes grew differentially in response to leaf extracts from the three species (although some studies^{7,8} on other tropical systems found little evidence for host preference). The degree of host specificity is a critical parameter for estimating total endophyte biodiversity based on isolation of fungi from individual tree species.

The most dramatic result occurred when Arnold *et al.* inoculated sterile cacao seedlings with several groups of endophyte, and then exposed the seedlings to a species of *Phytophthora*. This is a common pathogen that causes black pod disease and is related to the agent of Irish potato blight (*Phytophthora infestans*). Endophyte-free control leaves died in greater numbers than those inoculated with fungus, and the control leaves that did survive suffered much more damage (Fig. 2). The relative advantage of endophyte inoculation was higher in older than in younger leaves. Overall, endophyte-protected leaves suffered less than half the damage seen in control leaves (Fig. 1, inset).

But how does the fungal endophyte kill

or inhibit the growth of *Phytophthora*? One possibility is that the endophytes simply occupy the space that might be taken by the pathogen, or competitively consume resources; alternatively, direct antagonism might be involved. Furthermore, does inhibition result primarily from the action of a particular endophyte or do a variety of them act additively or synergistically? Would the results hold if endophytes and *Phytophthora* were inoculated simultaneously, or if the pathogen was inoculated first? In native

habitats, where endophytes and pathogens occur together, do heavily diseased plants host less diverse fungal communities?

Whatever the answers to these questions, Arnold and colleagues' findings¹ point to biological control as a promising way of tackling disease in cacao trees. This could be achieved by spraying the trees with fungal spores, or perhaps more realistically by growing them alongside other trees that could serve as sources of fungal inoculum. The results also clearly demonstrate that plant–endophyte mutualistic interactions are not restricted to cool-season grasses, and they explain why host plants don't defend themselves against infection by the fungi concerned. Pathogens and endophytes are common in all plants, so carrying out similar experiments in a range of tropical and temperate species should prove highly rewarding.

Keith Clay is in the Department of Biology, Indiana University, Bloomington, Indiana 47405, USA.

e-mail: clay@indiana.edu

1. Arnold, A. E. *et al.* *Proc. Natl Acad. Sci. USA* **100**, 15649–15654 (2003).
2. Carroll, G. C. *Ecology* **69**, 2–9 (1988).
3. Arnold, A. E., Maynard, Z., Gilbert, G. S., Coley, P. D. & Kursar, T. A. *Ecol. Lett.* **3**, 267–274 (2000).
4. Clay, K. & Schardl, C. L. *Am. Nat.* **160**, S99–S127 (2002).
5. Wilkinson, D. M. & Sherratt, T. N. *Oikos* **92**, 377–384 (2001).
6. Frohlich, J. & Hyde, K. D. *Biodiv. Conserv.* **8**, 977–1004 (1999).
7. Suryanarayanan, T. S., Venkatesan, G. & Murali, T. S. *Curr. Sci.* **85**, 489–493 (2003).
8. Cannon, P. F. & Simmons, C. M. *Mycologia* **94**, 210–220 (2002).

Nanotechnology

How does a nanofibre grow?

Pulickel M. Ajayan

Decades of research have failed to decipher the atomic-scale mechanism by which carbon nanofibres grow out of vapour. High-resolution microscopy shows that the carbon atoms have a bumpy ride.

In his historic monograph *On Growth and Form*¹, D'Arcy Thompson wrote that the form of an object is a diagram of its growth forces. Often, however, this relationship between growth and form is too complex to determine: the growth of filamentous carbon in the vapour phase is a case in point. Despite a large body of literature on the subject, replete with fascinating and varied examples, no definitive model for the growth of carbon nanofibres has evolved, owing to a lack of consistent experimental data^{2,3}. Considering the substantial impact that these materials are likely to have on technology⁴, it seems imperative that their growth mechanisms be understood, so that nanofibres can be manufactured with well-defined characteristics. A long-awaited solution to the mystery of nanofibre growth is presented

by Helveg *et al.*⁵ on page 426 of this issue.

Nanoscale carbon fibres are grown through the interaction of metal-catalyst nanoparticles with hydrocarbon vapour at high temperature. The hydrocarbon molecules dissociate at the interface between catalyst and vapour, and carbon atoms precipitate into a graphite trail in the shape of a cylindrical, multi-walled nanofibre (Fig. 1). Quite how the nanofibre forms is unknown. The catalyst particle might stay at the growing end of the nanofibre (called tip growth), or it might sit at the starting end (base growth)⁶. The state of the particle itself (such as its structure or shape) during the growth process is also unknown, but the particle size and fibre diameter are similar.

Experimentally, it has proved difficult to track the dynamics of this high-temperature catalytic reaction with spatial and temporal

resolution sufficient to observe the growth of the nanofibre directly at the atomic scale. The images presented by Helveg *et al.*⁵, capturing the early stages of nanofibre growth, are clearly the result of some impressive developments in high-resolution transmission electron microscopy *in situ*⁷. These authors have managed to record, in real time, the catalytic reaction at about 500 °C between methane and supported nickel nanoparticles (5–20 nm in diameter) inside a high-resolution transmission electron microscope. Analysed frame by frame, the images reveal the events that lead to nanofibre growth.

It seems that the catalyst particle moves with the growing nanofibre, following the tip-growth prescription. What is surprising is that the nanofibre growth is promoted by abrupt shape changes in the catalyst particle itself, which are in turn driven by the reaction between the catalyst and the vapour. These shape changes, between spherical and elongated, are sudden and repeated. In its elongated form the particle serves as a template, facilitating the formation of aligned graphite layers as carbon atoms diffuse across its surface. A closer look at the images reveals that the nucleation and growth of these graphite layers occur at 'bumps' on the nickel nanoparticle — single-atom step-edges that develop and then disappear continuously. This step-edge mechanism is also backed up by theoretical calculations.

Throughout the growth process, the nickel catalyst remains crystalline, bounded by the low-energy facets of the nanoparticle surface. The shape changes are driven by shifts in the balance of surface energy as graphite layers repeatedly cover then expose facets of the nickel surface. It is crucial that part of the nickel surface remains in direct contact with the surrounding vapour: when the particle is fully encapsulated by graphite layers, nanofibre growth stops because no further reaction between metal and vapour occurs.

These experiments by Helveg *et al.*⁵ show the importance of shape transformations in small particles when they are used as nanoscale growth catalysts. Indeed, the reaction-induced reshaping of metal particles seen here has certain parallels with the inherent structural instability of metal nanoparticles⁸. The diagram of forces that determines the final form of a nanofibre is closely linked to the drastic deformations of particle shape. As the particle gets larger, shape fluctuations become energetically unfavourable, and this could explain why large particles are inefficient catalysts for nanofibre growth⁹.

A word of caution is needed, however: the mechanism proposed by Helveg *et al.*⁵ may only be relevant in specific cases; it does not easily extend to smaller carbon nanotubes, in which the carbon lattice is far more ordered than in nanofibres and for which the catalyst particle is not always seen attached

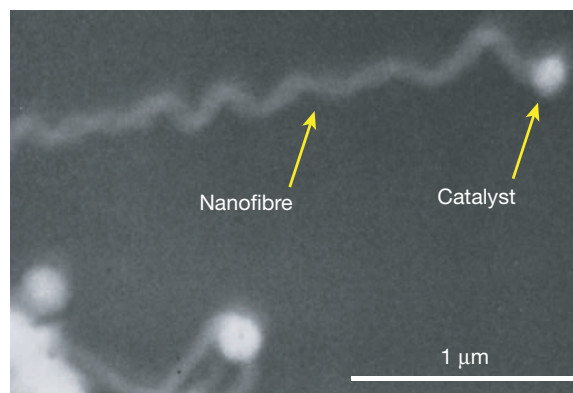


Figure 1 Growth of a nanofibre. At high temperatures, and in the presence of a catalyst, vapour-phase hydrocarbon molecules dissociate. The molecule's carbon atoms form into nanofibres, as shown in this scanning-electron-microscope image¹⁰. Helveg and colleagues' high-resolution study⁵ of the process now reveals the atomic-scale mechanism.

to the nanotube structure⁶. So, even with this new insight, the mystery of the growth of nanoscale filamentous carbon is not completely solved. But this study has overcome some major experimental stumbling blocks to provide the first direct glimpses of nanofibres as they grow. It heralds a new beginning in nanoscale real-time growth observations that should lead to better control over fibre synthesis for nanotechnology. ■

Pulickel M. Ajayan is in the Department of Materials Science and Engineering, Rensselaer Polytechnic Institute, Troy, New York 12180-3590, USA.

e-mail: ajayan@rpi.edu

1. Thompson, D.A. *On Growth and Form* (Cambridge Univ. Press, 1917).
2. Baker, R. T. K. & Harris, P. S. *Chem. Phys. Carbon* **14**, 83–165 (1978).
3. Tibbets, G. G. in *Carbon Filaments and Nanotubes: Common Origins, Differing Applications* (eds Biro, L. P. *et al.*) 63–73 (Kluwer, Dordrecht, 2001).
4. Terrones, M. *Annu. Rev. Mater. Res.* **33**, 419–501 (2003).
5. Helveg, S. *et al. Nature* **427**, 426–429 (2004).
6. Charlier, J. C. & Iijima, S. in *Topics in Applied Physics* Vol. 80 (eds Dresselhaus, M. S. *et al.*) 55–80 (Springer, Berlin, 2001).
7. Hansen, P. *et al. Science* **295**, 2053–2055 (2002).
8. Ajayan, P. M. & Marks, L. D. *Phys. Rev. Lett.* **63**, 279–282 (1989).
9. Nerushev, O. A., Dittmar, S., Morjan, R.-E., Rohmund, F. & Campbell, E. E. B. *J. Appl. Phys.* **7**, 4185–4190 (2003).
10. Banhart, F. *et al. Int. J. Mod. Phys. B* **15**, 4037–4069 (2001).

Developmental biology

Tail of decay

Alexander F. Schier

Humans and other vertebrates develop in a head-to-tail sequence. A mechanism that is based on a gradual decay of RNA appears to contribute to this process.

The thousands of different cell types in the vertebrate body are arranged in a complex but reproducible pattern. The underlying blueprint develops in the embryo, when initially naive cells mature and specialize according to their position. In recent years, it has become clear that positional information is often provided by the graded distribution of signals^{1,2}. Depending on their location in the gradient, cells are exposed to different concentrations of the signal and differentiate in a dose-dependent way. But how do these gradients form? On page 419 of this issue, Dubrulle and Pourquie³ describe one mechanism. In the growing embryo, cells at the tip of the tail produce a messenger RNA (mRNA) molecule that codes for a protein signal; as cells move away from the tip, the levels of mRNA gradually decay. The resulting gradient is thought to regulate the maturation of cells.

The use of gradients and dosage-dependent responses is widespread in development, from fly embryos to the mammalian nervous system. A gradient has also been

implied in the head-to-tail development of vertebrates. During early development, embryonic structures are progressively laid down from head (anterior) to tail (posterior). This process can be illustrated by the formation of the skeleton. The vertebrae are arranged in a repetitive anterior-to-posterior pattern. They each emerge from small blocks of tissue called somites (Fig. 1a, overleaf). The first somite forms most anteriorly, and then one somite after the other is added in a head-to-tail sequence⁴. A somite forms through maturation of the presomitic mesoderm — a region of immature cells that lie posteriorly to the last-formed somite. The presomitic mesoderm in turn derives from cells that are released from a more posterior growth zone called the tail bud. So, during their maturation, cells move in an assembly line from the tail bud to the presomitic mesoderm and then to the somites.

Previous studies have shown that this maturation is controlled by a signalling molecule that belongs to the fibroblast growth factor (FGF) family of proteins⁵. It appears

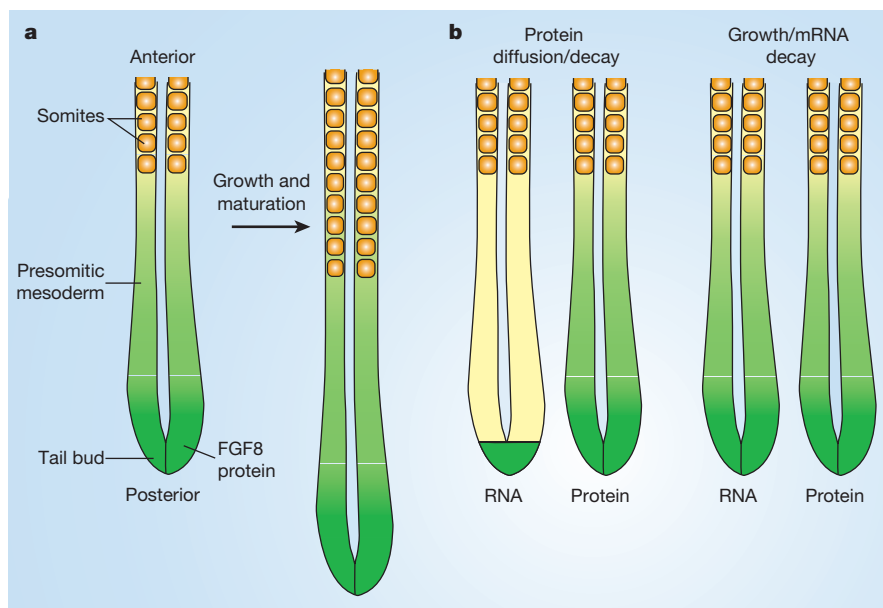


Figure 1 Head-to-tail somite development in vertebrates. **a**, Three main regions can be distinguished in the embryo: the tail bud, presomitic mesoderm and somites. During maturation, cells originate in the tail bud, then move to become presomitic mesoderm and finally form a somite. The FGF8 protein gradient determines where cells form somites; high levels keep cells in an immature state and lower levels allow maturation. As the embryo grows, the FGF8 gradient regresses in an anterior-to-posterior direction, resulting in the anterior-to-posterior direction of maturation. **b**, Models for gradient formation. In the 'protein diffusion/decay' model, FGF8 mRNA and protein are made only in the tail bud. The protein is secreted, diffuses away and gradually decays, resulting in a gradient. In the 'growth/mRNA decay' model proposed by Dubrulle and Pourquié³, again only tail-bud cells make FGF8 mRNA. But in this case, as the cells leave the tail bud the mRNA gradually decays, resulting in an FGF8 mRNA gradient that is translated into a protein gradient. As Dubrulle and Pourquié describe, the levels of mRNA in neighbouring cells are not necessarily the same, because cells often change their position; however, the overall distribution of FGF8 mRNA and protein is graded.

that high levels of FGF8 keep cells in the immature 'tail bud' and 'presomitic mesoderm' states (Fig. 1a). At a certain threshold, however, the levels of FGF8 become too low to be effective, and maturation and somite formation are triggered. In support of this model, Dubrulle and Pourquié³ report that, in chick embryos, the FGF8 protein forms a posterior-to-anterior gradient that is translated into graded activation of the downstream protein Akt. Hence, an inverse FGF signalling gradient seems to underlie the head-to-tail gradient of cell maturation. As the embryo grows and the tail bud extends, the FGF8 gradient also regresses in an anterior-to-posterior direction, setting off a head-to-tail wave of maturation.

But how is the FGF8 gradient formed? Previous work had suggested that the most common mechanism of gradient formation in development is the locally restricted expression of a gene encoding a secreted protein signal¹ (the gene is transcribed into mRNA, and the mRNA is translated into protein). According to this model, the cells that produce the protein act as a source and release the signal to surrounding cells, where it gradually decays, resulting in the formation of a gradient. If this 'protein diffusion/decay' model were to apply to somite

formation, FGF8 mRNA and protein would be produced locally in the tail bud (Fig. 1b). FGF8 protein would then move away from this source, decay and form a gradient. Although this mechanism is likely to contribute to FGF gradient formation, the paper by Dubrulle and Pourquié³ suggests an additional, quite different model.

It turns out that there is not only an FGF8 protein gradient, but also an FGF8 mRNA gradient (Fig. 1b). Tail-bud cells have high levels of FGF8 mRNA, whereas more anterior cells have lower levels. Hence, FGF8 mRNA is also translated into protein in more anterior regions — not only, as the protein diffusion/decay model postulates, in the tail bud. Apparently, this translation of graded FGF8 mRNA results in a graded FGF8 protein distribution.

The question then becomes, how does this mRNA gradient form? In an elegant experiment, Dubrulle and Pourquié use intronic sequences (segments of a precursor mRNA that do not end up in the mature molecule) to detect where the FGF8 gene is being transcribed. Using chick embryos, they find that FGF8 precursor mRNA is made only in the tail bud, not in all regions where the mature mRNA is present. In addition, blocking gene transcription in the anterior region

does not eliminate FGF8 mRNA from that region. These results suggest that only cells in the tail bud actively transcribe the FGF8 gene. When these cells leave the tail bud and form the presomitic mesoderm and somites, they cease to make FGF8 mRNA and their supply begins to degrade. This results in an FGF8 mRNA gradient.

This 'growth/mRNA decay' model is quite different from the protein diffusion/decay mechanism. Instead of having a protein that is generated locally, moves away from the source and decays over time, we are now dealing with cells that contain locally generated mRNA and move away from the tail bud, while the mRNA gradually decays. In both cases, however, local production, movement and decay result in the formation of a gradient.

The finding of an alternative mechanism for generating gradients is exciting. But what would be the advantage of this system compared with the protein-diffusion mechanism? There are several possibilities. The FGF8 protein might be too unstable to produce a long-range gradient (this could be particularly true for the large and rapidly growing chick embryo). In this case, the growth/mRNA decay mechanism would ensure the presence of FGF8 protein far away from the tail bud. In addition, the slope of a protein gradient formed by diffusion could be very steep. In contrast, the mRNA gradient could result in a more shallow protein gradient, potentially allowing more fine-grained patterning.

These considerations raise one remaining question: it is not yet clear whether the growth/mRNA decay mechanism is indeed required for embryonic patterning. It could be that the local production of FGF8 protein in the tail bud is enough to generate a gradient of FGF8 protein that allows normal development. Blocking FGF8 mRNA translation in the tail bud or more anterior regions could address this issue.

How common is this mechanism of gradient formation? FGF8 might represent the tip of an iceberg of genes that are regulated in this fashion. And is this mechanism widely involved in the patterning of tissues that undergo polarized growth? Because the approaches used by Dubrulle and Pourquié³ can be easily applied to other genes and tissues, answers to these questions can be expected soon.

Alexander F. Schier is at the Skirball Institute of Biomolecular Medicine and in the Department of Cell Biology, New York University School of Medicine, 540 First Avenue, New York, New York 10016, USA.

e-mail: schier@saturn.med.nyu.edu

1. Vincent, J. P. & Dubois, L. *Dev. Cell* **3**, 615–623 (2002).
2. Gurdon, J. B. & Bourillot, P. Y. *Nature* **413**, 797–803 (2001).
3. Dubrulle, J. & Pourquié, O. *Nature* **427**, 419–422 (2004).
4. Pourquié, O. *Science* **301**, 328–330 (2003).
5. Dubrulle, J., McGrew, M. J. & Pourquié, O. *Cell* **106**, 219–232 (2001).

Earthquake science

Faults greased at high speed

Chris Marone

The dynamics of the tectonic faults that produce earthquakes remain puzzling. An inference from laboratory experiments could help: at high rates of slip, friction at the interface may fall dramatically.

For several decades now, geophysicists have been trying to understand why the energy budget for tectonic faulting does not seem to add up. The problem is that faults appear to be more slippery — less constrained by friction — than has been predicted by laboratory and theoretical work. The measurements of rock friction, described on page 436 of this issue by Di Toro and colleagues¹, may put things on firmer ground. They show that friction of quartz-rich rock is indeed high at low slip rates, consistent with previous studies, but that under certain conditions it drops dramatically as slip velocity approaches a few millimetres per second.

Conventional wisdom is that missing expenses are the cause of the imbalance in the faulting budget. On the income side of this budget are the driving forces of plate tectonics and the elastic energy stored in Earth's crust. Laboratory measurements of fault-zone friction indicate that the frictional stress during faulting near Earth's surface should be of the order of 50–100 megapascals. This implies that substantial frictional heat is produced during faulting, because the other main energy expenses — radiation of seismic waves, and the creation of surface area from the production and comminution of 'wear material' — are thought to account for only a small fraction of energy dissipation. The problem is that the expected frictional heat is missing^{2,3}.

To study the strength of frictional contact between rock surfaces, Di Toro and colleagues used an apparatus that applies rotary shear to the samples. The apparatus allows only comparatively small unidirectional movement, so the authors sheared samples back and forth to achieve the large net displacements that occur at earthquake faults. As in previous experiments in geophysical rock mechanics, they sheared samples under the high stresses expected to apply at tectonic faults.

Consistent with existing data, Di Toro *et al.* found that the coefficient of friction was 0.6–0.7 at low sliding velocities (up to 1 mm s⁻¹). However, their experiments show that the coefficient for novaculite — a rock composed of silicon dioxide (quartz) — decreases dramatically to values as low as 0.2 when the shearing velocity exceeds 1–10 mm s⁻¹. This effect is transient. On returning to lower sliding velocity, the coefficient of friction returns to high values. Identical experiments on samples of granite did not show the same reduced friction (or

'weakening') at high speed. So the authors infer that the weakening mechanism is related to the formation of a thin layer of silica gel, which acts as grease between the surfaces.

It seems likely that, in addition to the presence of quartz, a key factor in these experiments is the low rate of heat production. Previous laboratory studies of high-speed rock friction did not find appreciable weakening⁴ and reported significant shear melting, possibly because they involved greater acceleration and faster shear rates than those studied by Di Toro and colleagues. Further work is needed to understand the weakening mechanism discovered by Di Toro *et al.*, but their findings may provide a helpful clue in balancing the faulting budget. As applied to tectonic faults, the results indicate that frictional stress and shear heating may be much lower than expected — implying that faulty assessments of income, rather than of expenses, are at the root of the energy budget imbalance.

There are several reasons for caution in attempting to extrapolate the new results to tectonic faults. Foremost among these are the roughness of faults and the thick accumulations of 'fault gouge' — granular wear material — within fault zones. Samples used in laboratory friction experiments are very smooth compared with natural faults, and it is not clear if a thin layer of silica gel could effectively lubricate a rough, gouge-filled fault. Laboratory experiments⁵ on quartz gouge do not exhibit the extreme frictional weakening seen by Di Toro *et al.* Furthermore, fault rocks and gouge tend to be composed of several minerals, which would argue against a weakening mechanism that operated only on quartz-rich rocks.

Nonetheless, several lines of evidence seem to point to a high-speed, dynamic weakening of fault zones⁶. The threshold weakening velocity reported by Di Toro *et al.* would help in relating this dynamic weakening to seismic data. Previous analyses of frictional weakening^{7,8} have been predicated on the expectation that considerable weakening would occur only for large earthquakes of magnitude 5 or greater. However, seismic-stress reduction — that is, the drop in frictional stress caused by earthquakes — is remarkably consistent over at least six orders of magnitude in fault dimension, and earthquake frequency–magnitude distributions obey a self-similar power-law relation over this same range. These observations are good evidence that the physics of



100 YEARS AGO

An interesting paper on a familiar subject, the relation of temperature to the keeping property of milk, has reached us from Storrs, Connecticut. The view of the writer, Mr. H. W. Conn, the well-known dairy bacteriologist, is that the keeping of milk is more a matter of temperature than of cleanliness. He points out that at 50° F. milk may not curdle for two weeks, whereas at 70° F. it may keep but forty-eight hours, and at 95° F. but eighteen hours. This curdling is due to the action of bacteria, and the effect of temperature on their multiplication is surprising. Thus at 50° the ordinary milk organisms increase about 5-fold in twenty-four hours, but at 70° they may multiply 750-fold in the same time. From *Nature* 28 January 1904.

50 YEARS AGO

At a conference called by the Federation of British Industries in London, on January 14, to discuss the shortage of science teachers in schools, the president, Sir Harry Pilkington, said that the national interest demands that enough science teachers of quality, for all types of educational institution, shall teach and train students in the next few years, so that our scientific and industrial leadership of the world over the next generation may resist challenges... Of the good academic men with fine personal qualities — of whom industry, research and the universities and schools must all have some — there are not enough to go round... The answer to the shortage of graduate teachers in science is that "we must make sure we are using well all the scientists we have in research and in industry and that we must somehow again make schoolmastering attractive to the young man of spirit, imagination and ambition". How is this to be done? More and more people are coming to think that the benefits of salary and other differentials must go to good men who actually do the teaching as opposed to good organizers and administrators in the schools... There is considerable evidence, culminating in the latest report of the Advisory Council on Scientific Policy, to show that Great Britain is already behind some of its important industrial competitors in the application of science. To meet this situation we need not only a restoration of our former standards of science teaching but also a considerable advance on these standards.

From *Nature* 30 January 1954.

earthquake rupture is the same for small and large earthquakes.

The work of Di Toro *et al.* implies that dynamic weakening does not have a threshold determined by earthquake magnitude: in their experiments the onset of weakening occurs at only a few millimetres per second, which would be reached even for very small earthquakes. In the context of this model, all earthquakes would experience considerable dynamic weakening and there would be no break in earthquake scaling relations. Furthermore, because shear heating would be negligible, dissipation of seismic energy would be dominated by radiation of seismic waves and creation of surface area.

The new results are a step forward in understanding rock friction and how it may change as faults accelerate from slip rates associated with creep and plate tectonic motion (millimetres per year) to seismic slip rates of metres per second. It will be necessary to find out if the weakening mech-

anism described applies to rough surfaces and can be extended to tectonic faults. If it can, we would have at least a partial solution to the problematic energy budget for faulting. But faults that experience substantial aseismic creep and do not generate earthquakes would still be expected to have high frictional strength and significant shear heating. ■

Chris Marone is in the Department of Geosciences, 536 Deike Building, Pennsylvania State University, University Park, Pennsylvania 16802, USA.

e-mail: cjm@geosc.psu.edu

1. Di Toro, G., Goldsby, D. L. & Tullis, T. E. *Nature* **427**, 436–439 (2004).
2. Saffer, D. M., Bekins, B. A. & Hickman, S. J. *Geophys. Res.* **108**, doi:10.1029/2002JB001849 (2003).
3. Scholz, C. H. *Geology* **28**, 163–166 (2000).
4. Spray, J. G. *J. Geophys. Res.* **98**, 8053–8068 (1993).
5. Mair, K. & Marone, C. J. *Geophys. Res.* **104**, 28899–28914 (1999).
6. Lapusta, N. & Rice, J. R. *Eos* **84**, S51B-02 (2003).
7. Heaton, T. *Phys. Earth Planet. Inter.* **64**, 1–20 (1990).
8. Kanamori, H. & Heaton, T. H. in *Geocomplexity and the Physics of Earthquakes* (eds Rundle, J., Turcotte, D. L. & Kein, W.) AGU Monogr. 120, 147–155 (AGU, Washington DC, 2000).

Materials science

A natural solution to corrosion?

Stuart Lyon

Corrosion damage can be reduced if inhibitor molecules are introduced into a metal's environment. As inhibitors may themselves be noxious, the inhibitory properties of natural amino acids are now under scrutiny.

It is one of those peculiar aspects of human nature that certain topics are not to be discussed in polite company. If such discussion becomes unavoidable, we disguise the true nature of the activity by euphemism. So, a rat catcher becomes a 'rodent control operative' and a toilet cleaner might be styled 'sanitary inspector'. Of course, being human, we scientists and engineers also aim to improve the image of our fields of endeavour. Hence, studies in pollution become 'environmental protection' and hot-dip galvanized steel is now 'surface engineered'. Sadly, some topics appear forever irredeemable.

Corrosion is both one of those topics that no one really wants to talk about, and one that seems to have no acceptable alternative description. It is also a subject that touches all our lives. In industrialized economies, material degradation through corrosion costs at least 2–3% of gross national product per annum. However, up to a third of these costs might be saved through the implementation of advanced techniques for controlling corrosion. One such technique, using natural amino acids, is now pursued by Ashassi-Sorkhabi and colleagues¹, in work published in *Applied Surface Science*.

In practice, corrosion may be reduced by contriving to modify the chemistry of the material surface, or of its environment, to slow the rate of any reaction between them.

For example, environmental modification may involve the addition of small quantities of chemical inhibitors, which profoundly affect the kinetics of the individual electrochemical reactions that together comprise the corrosion process. Corrosion inhibitors are widely used, for example in the water treatment of boilers, and in commercial and domestic heating and cooling systems, where conventional formulations often include species (such as nitrite or benzoate) that have significant toxicity. As a result, proper disposal is awkward and comes at a substantial cost.

Another complication arises in the form of the mineral acids that are typically used in industrial processes for cleaning and descaling metals, and also to acidify oil wells — here, a strong acid, such as hydrochloric acid, is injected at intervals into a well to aid the break-up of rock strata and ease the extraction of oil. The presence of acid will obviously speed up the corrosion of the carbon-steel well-casings and pipes of the drilling rigs (Fig. 1), although inhibitors can be added to control it. But, once again, large quantities of used, contaminated, inhibitor-treated acid cannot be simply neutralized and dumped: inhibitors such as phosphonates, alkynes and quaternary ammonium salts are significant biohazards and pollutants in their own right.

It was therefore with some interest that I read Ashassi-Sorkhabi and colleagues' paper¹,

in which they consider the use of the simple amino acids alanine, glycine and leucine as corrosion inhibitors for carbon steel in hydrochloric acid. In acid inhibition, effective inhibitors generally form a monolayer (or perhaps multilayers) on a metallic surface, which is oxidized to some degree, depending on the pH of the environment; electron-donating groups such as N, O, S and P on an inhibitor molecule can often enhance strong adsorption. Because the inhibitor layer interferes with electron transfer and excludes the active acid from the metal surface, the rate of corrosion is reduced — although this simplistic mechanism is by no means the whole story.

In studying the effect of their chosen amino acids, Ashassi-Sorkhabi and colleagues have used a standard electrochemical method, called potentiodynamic polarization, to work out the surface coverage of each species on the metal. The electrochemical potential of the metal in the environment is brought to some negative value, then swept to a more positive value and the current is measured simultaneously. Changes in the current flow as the environment changes (that is, as the concentration of amino acid in the hydrochloric acid is varied) provide information about the reaction kinetics at the acid-metal interface, and hence about the overall rate of corrosion.

The three amino acids prove to be quite functional — although not outstanding — corrosion inhibitors. At concentrations of 0.1 and 0.01 moles in 0.1-molar acid, their presence reduces the corrosion rate by between 50% and 90%, which implies at least a doubling of component lifetime. And, of course, the disposal of a naturally occurring amino acid is considerably less of a problem than for more conventional inhibitors.



Figure 1 Acid attack: drill pipes on oil rigs suffer corrosion.

GETTY IMAGES

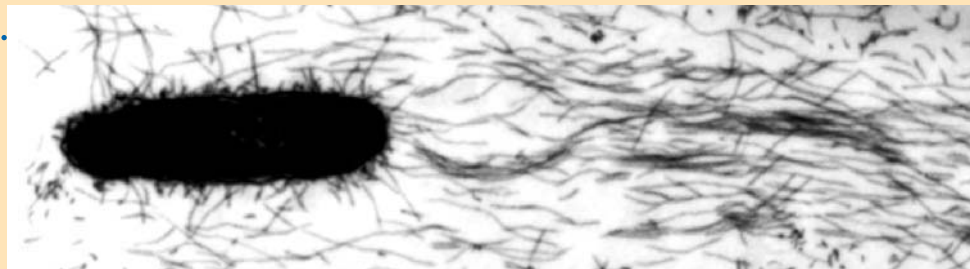
Cell biology

Pathogen propulsion

Bacterial pathogens have a cunning way of moving about inside the cells they infect — they harness components of the host's cytoskeletal machinery, in particular, the protein actin.

One such pathogen, *Rickettsia conorii*, is transmitted to humans through tick bites, and causes Mediterranean spotted fever. This bacterium assembles an elaborate 'tail' made of actin filaments (pictured) to propel itself through the cytoplasm of the infected host cell and to invade neighbouring cells. *Rickettsia* tails consist of parallel, unbranched filaments that closely resemble those present in thread-like cellular projections called filopodia.

Elsewhere in this issue, Pascale



Cossart and colleagues shed light on the molecular mechanism that generates *Rickettsia* tails (E. Gouin *et al. Nature* **427**, 457–461; 2004). The authors had previously identified a *Rickettsia* protein, RickA, that is structurally similar to members of the WASP family found in higher organisms. WASPs are potent activators of the Arp2/3 complex, which nucleates actin filaments.

In the latest development, Cossart's group has found that RickA occurs on the bacterial

surface, where actin filaments are generated. Using an *in vitro* actin polymerization assay, they go on to show that although RickA on its own cannot form actin filaments, it can activate the Arp2/3 complex — thereby stimulating actin polymerization. They extend these studies to show that *Rickettsia* uses the Arp2/3 complex *in vivo* to form actin tails, and that RickA can induce the formation of filopodium-like projections when it is targeted to the plasma membrane. So future work on *Rickettsia* motility might

help to clarify the poorly understood process of filopodium formation.

Rickettsia actin tails are very different to those created by another pathogen, *Listeria monocytogenes*, which consist of shorter, highly branched arrays of actin filaments. Nevertheless, the *Listeria* tails are also formed by the Arp2/3 complex, which the bacterium recruits from the host cell using the surface protein ActA. So pathogens may have evolved different ways to induce actin polymerization using the same effector, the Arp2/3 complex.

Deepa Nath

Ashassi-Sorkhabi and colleagues' work got me wondering whether there were any other natural compounds that act as inhibitors. A quick bibliographic search revealed a long history of research on the anti-corrosion properties of naturally occurring materials: extracts of prickly pear², henna³, rosemary⁴ and honey⁵ have all been cited as providing a greater than 50% reduction in the corrosion rate of carbon steels, aluminium alloys and copper alloys, in chloride media of up to one-molar concentrations. And historically, tannins⁶ were used in the steam age to reduce boiler corrosion.

Particularly fascinating, however, is recent work using organisms (bacillae or fungi) that have been genetically modified to secrete corrosion-inhibiting species, such as polyphosphate⁷. Normally, biofilms tend to accelerate corrosion, but inoculation with these 'smart' organisms can, in appropriate environments, substantially reduce it. Such technology potentially offers immense benefits for environmentally friendly corrosion control. I look forward to the day when we might all be advised to add yoghurt to our central heating systems! Corrosion might then cease to be the hidden enemy no one wants to discuss. ■

Stuart Lyon is in the Corrosion and Protection Centre, University of Manchester Institute of Science and Technology, Manchester M60 1QD, UK.

e-mail: stuart.lyon@umist.ac.uk

1. Ashassi-Sorkhabi, H., Majidi, M. R. & Seydidi, K. *Appl. Surf. Sci.* doi:10.1016/j.apsusc.2003.10.007 (2003).
2. El-Etre, A. Y. *Corros. Sci.* **45**, 2485–2495 (2003).
3. Chetouani, A. & Hammouti, B. *Bull. Electrochem.* **19**, 23–25 (2003).
4. Kliskic, M. *et al. J. Appl. Electrochem.* **30**, 823–830 (2000).
5. El-Etre, A. Y. & Abdallah, M. *Corros. Sci.* **42**, 731–738 (2000).
6. Matamala, G., Smeltzer, W. & Droguett, G. *Corros. Sci.* **42**, 1351–1362 (2000).
7. Ornek, D. *et al. Appl. Microbiol. Biotechnol.* **58**, 651–657 (2002).

Molecular motors

Turning the ATP motor

Richard L. Cross

A long-standing question regarding ATP synthase — a cellular energy-generator — has been which direction it spins in when generating ATP. Some elegant experiments have revealed the answer.

The ATP synthase not only lays claim to being nature's smallest rotary motor, but also has an extremely important role in providing most of the chemical energy that aerobic and photosynthetic organisms need to stay alive. On page 465 of this issue, Itoh and colleagues¹ describe how they used electromagnets to force this motor to rotate and generate chemical energy (adenosine triphosphate, ATP). In the process, they determined the direction of rotation during ATP synthesis.

ATP synthase is composed of two linked multi-subunit complexes, called F_0 and F_1 . F_0 is embedded in cellular membranes and conducts protons, whereas F_1 is a peripheral complex and contains the catalytic sites. Together they couple the flow of protons down an electrochemical gradient to the synthesis of ATP from ADP (adenosine diphosphate) and inorganic phosphate.

The major — and initially controversial — features of this process were first recognized by Paul Boyer, who developed the concept of the 'rotary binding-change' mechanism. The first surprise came from his discovery that the energy derived from proton transport is not used to promote the

synthesis of ATP at the catalytic sites on F_1 . Instead, ATP forms spontaneously from tightly bound ADP and phosphate (Fig. 1a, step 2, overleaf); the energy instead drives the subsequent release of ATP from that site². Also associated with this dissociation step (Fig. 1a, step 1) is an increase in the affinity of adjacent catalytic sites for ADP and phosphate³. A second surprise was the recognition that the coupling process probably involves subunit rotation⁴.

In the binding-change model, then, F_0 and F_1 function as a pair of rotary motors linked by a central rotor and a peripheral stator (Fig. 1b). Rotation of the ring of c -subunits in F_0 is proposed to allow protons to be carried between two channels, formed by the a -subunit, that connect to opposite sides of the membrane. Because of the existence of the stator^{5,6}, and because the γ - and ϵ -subunits of the central rotor are firmly attached to the top of the c -ring⁷, rotation of the c -ring forces the γ -subunit to rotate in the centre of F_1 . This in turn drives net ATP synthesis by inducing cyclical conformational changes in the surrounding subunits, thereby altering binding affinities at the catalytic sites (hence 'binding-change' model; Fig. 1a, step 1). Strikingly, under

conditions in which more energy is stored in the ATP pool than in the electrochemical proton gradient, the engine can carry out the opposite reaction, consuming (hydrolysing) ATP to rotate F_0 and pump protons up a concentration gradient.

The rotary aspect of this theory remained a popular but speculative idea for years, until compelling evidence was provided by structural⁸, biochemical⁹ and spectral¹⁰ studies. Then, in a dramatic visual demonstration, a fluorescent actin filament attached to one end of the γ -subunit of immobilized F_1 was seen by fluorescence microscopy to undergo multiple unidirectional rotations during ATP hydrolysis¹¹. The direction of rotation, as viewed from the position of F_0 , was anticlockwise. Subsequent biochemical experiments¹² established that energy-driven rotation of the γ -subunit occurs during ATP synthesis as well. But in this case the direction of rotation could not be determined by the approach used. So the question of whether or not the direction of rotation reverses when the direction of catalysis reverses has, until now, remained unanswered.

In this regard, it is interesting that the only other well-characterized biological rotary motor — the bacterial flagellar motor — has a transmission that, in response to regulatory signals, can shift the direction of rotation between forward and reverse without changing the direction of proton movement through the membrane. In contrast, the ATP synthase is smaller and simpler, and because of this it has long been thought that the direction of the coupled chemical events (ATP synthesis or hydrolysis) and proton transport (down or up a gradient) would indeed be reversed upon reversing the direction of subunit rotation. In the experiments presented by Itoh *et al.*¹, this prediction is at last confirmed.

The authors accomplished this feat by attaching magnetic beads to the γ -subunits on isolated F_1 complexes, which were immobilized by fixing their tops to a glass surface. The authors then used electromagnets to rotate the beads (substituting for proton transport), which in turn forced the γ -subunits to rotate. ADP and inorganic phosphate were included in the experiment, as well as an assay for monitoring the concentration of ATP. Itoh *et al.* found that rotating the γ -subunit in the anticlockwise direction — previously shown to accompany ATP hydrolysis — caused a decrease in ATP levels. But rotating the γ -subunit in a clockwise direction caused an increase in ATP. The results establish a hard-wired reversibility in the F_0F_1 rotary motor.

An interesting feature of Itoh and colleagues' experiment is that it works only at low rotational speeds, that is, roughly 5% of the maximal ATP-driven rate. As the speed increases, bead rotation becomes uncoupled from the magnet. This is not surprising

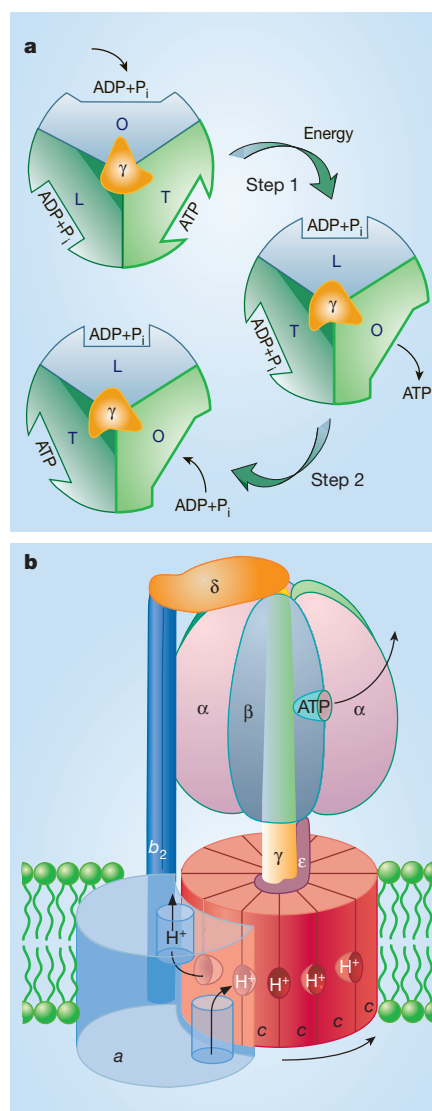


Figure 1 The binding-change model for F_0F_1 ATP synthase. **a**, Looking up at F_1 from the membrane. Each blue or green area represents a pair of α - and β -subunits, in which the catalytic sites are interfacial but mostly on the β -subunit. In step 1, the γ -subunit rotates through 120° , driving conformational changes in the three surrounding catalytic sites that alter their affinities (O, L or T, for open, loose or tight) for substrates and product. In step 2, ATP forms spontaneously from tightly bound ADP and inorganic phosphate (P_i). **b**, View from the side of F_0F_1 . The a -subunit contains two partial channels. For a proton to traverse the membrane, it must move through one channel to the centre, bind to one of the c -subunits and then be carried to the other partial channel by rotation of the c -ring. The c -subunits are anchored to the γ -subunit (part of the rotor), whereas the a -subunit is anchored through $b_2\delta$ (the stator) to the $\alpha_3\beta_3$ hexamer. Hence, rotation of the c -ring relative to the a -subunit in F_0 will drive the rotation of the γ -subunit relative to the $\alpha_3\beta_3$ hexamer in F_1 . New results^{1,15} show that the γ -subunit rotates in a clockwise direction when the engine generates ATP.

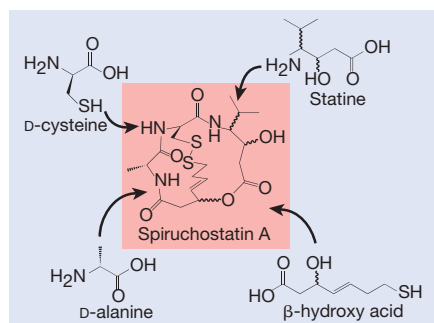
because the rotary torque is applied at a constant rate in the experiments, whereas under physiological conditions the γ -subunit can stop rotating for significant periods during substrate binding, product release, and catalysis^{13,14}. So it might have been predicted that, with increasing speed, a point would be reached at which the pauses limit bead rotation such that it falls out of synchronization with the rotating magnetic field.

As often happens with significant advances, another laboratory has independently reached the same conclusion about the mechanical properties of ATP synthase. Diez and colleagues¹⁵ report the use of fluorescence resonance energy transfer to measure the distance between a pair of probes attached to the b - (stator) and γ - (rotor) subunits. Three distinct distances are clearly evident, and correspond to the orientation of the γ -subunit when paused at 120° intervals (see Fig. 1a). As expected, the distance between the probes changes during catalytic turnover, in repeat sequences that progress at a rate equal to that of catalysis. Furthermore, the order of the sequence reverses in going from ATP hydrolysis (close to intermediate to distant) to ATP synthesis (distant to intermediate to close), again establishing a correspondence between the direction of catalysis and the direction of rotation. These studies^{1,15} not only further our understanding of an essential energy-transducing complex but also enhance the prospects for exploiting this rotary motor to develop nanodevices.

Richard L. Cross is in the Department of Biochemistry and Molecular Biology, SUNY Upstate Medical University, Syracuse, New York 13210, USA.

e-mail: crossr@upstate.edu

1. Itoh, H. *et al.* *Nature* **427**, 465–468 (2004).
2. Boyer, P. D., Cross, R. L. & Momsen, W. *Proc. Natl Acad. Sci. USA* **70**, 2837–2839 (1973).
3. Kayalar, C., Rosings, J. & Boyer, P. D. *J. Biol. Chem.* **252**, 2486–2491 (1977).
4. Boyer, P. D. & Kohlbrenner, W. E. in *Energy Coupling in Photosynthesis* (eds Selman, B. & Selman-Reiner, S.) 231–240 (Elsevier North-Holland, New York, 1981).
5. Boekema, E. J., Ubbink-Kok, T., Lolkema, J. S., Brissos, A. & Konings, W. N. *Proc. Natl Acad. Sci. USA* **94**, 14291–14293 (1997).
6. Wilkens, S., Zhou, J., Nakayama, R., Dunn, S. D. & Capaldi, R. A. *J. Mol. Biol.* **295**, 387–391 (2000).
7. Hermolin, J., Dmitriev, O. Y., Zhang, Y. & Fillingame, R. H. *J. Biol. Chem.* **274**, 17011–17016 (1999).
8. Abrahams, J. P., Leslie, A. G., Lutter, R. & Walker, J. E. *Nature* **370**, 621–628 (1994).
9. Duncan, T. M., Bulgin, V. V., Zhou, Y., Hutcheon, M. L. & Cross, R. L. *Proc. Natl Acad. Sci. USA* **92**, 10964–10968 (1995).
10. Sabbert, D., Engelbrecht, S. & Junge, S. *Nature* **381**, 623–625 (1996).
11. Noji, H., Yasuda, R., Yoshida, M. & Kinosita, K. Jr *Nature* **386**, 299–302 (1997).
12. Zhou, Y., Duncan, T. M. & Cross, R. L. *Proc. Natl Acad. Sci. USA* **94**, 10583–10587 (1997).
13. Sabbert, D. & Junge, W. *Proc. Natl Acad. Sci. USA* **94**, 2312–2317 (1997).
14. Adachi, K. *et al.* *Proc. Natl Acad. Sci. USA* **97**, 7243–7247 (2000).
15. Diez, M. *et al.* *Nature Struct. Mol. Biol.* **11**, 135–141 (2004).



Synthetic chemistry

Spiruchostatin from scratch

J. Am. Chem. Soc. doi:10.1021/ja039258q (2004)

Histone deacetylases are enzymes involved in regulating DNA transcription, and several molecules that inhibit their activity are being investigated as potential anticancer drugs. In particular, the natural product FK228, derived from *Chromobacterium violaceum*, is currently in phase II clinical trials.

To understand exactly how FK228 works in the body, and to discover whether nature's own anticancer agents can be improved on, chemists need a range of analogues of FK228 to work with. But FK228 has proved very resistant to any structural modifications. One solution is to build the molecule from scratch, using a method that is flexible enough to produce a variety of similar molecules with the same core structure.

Alexander Yurek-George and colleagues have developed the first total synthesis of spiruchostatin A, a close analogue of FK228 that was originally isolated from an extract of the bacterium *Pseudomonas*. At a crucial stage of the synthesis, the authors introduce a Nagao chiral auxiliary — a chemical that attacks only one side of the substrate molecule, preferentially forming a product that is easily separated from trace amounts of an unwanted mirror-image compound. This not only puts them on the road to obtaining the correct form of spiruchostatin A, but also activates a key intermediate so that the two halves of the final product come together with ease.

Mark Peplow

Cancer

Assisted suicide for tumours

Cancer Cell 5, 25–35 (2004)

Defective cells eliminate themselves by a process called apoptosis — tumours may arise if they fail to do so. Aaron D. Schimmer and colleagues have screened nearly one million compounds to find a handful that jump-start the apoptosis programme in a wide variety of tumour cells.

The effectors of the cell's death programme are the caspases. These cellular

knives cut specific proteins with precision to bring about the systematic collapse of a cell. In normal cells, caspases lie dormant and their activation indicates that the cell's end is near. Like sharp tools fitted with safety caps, caspases are physically kept in check by 'inhibitor of apoptosis' proteins — IAPs. But these safety caps may also help tumour cells to survive and defy chemotherapy. For example, one IAP, called XIAP, is present at high levels in various cancers and strongly inhibits caspase-3 and caspase-7.

Schimmer *et al.* found a few small compounds that could release XIAP from caspase-3 and unleash the latter's cutting activity. These compounds induced apoptosis of tumour cell lines and sensitized them to anticancer drugs. Moreover, they limited tumour growth in mice, showing surprisingly little toxicity to normal tissues. The authors suggest that in cancer cells caspases might no longer be dormant, so that the simple removal of IAPs allows apoptosis to proceed.

Marie-Thérèse Heemels

Microelectronics

Keep the beat

Phys. Rev. Lett. 92, 027201 (2004)

On-chip clocks for synchronizing the timing of electronic circuits could result from a device created by W. H. Rippard and colleagues. The group induced oscillations at microwave frequencies (about 5–40 gigahertz) in a structure built from stacked layers of metals and alloys each just a few nanometres thick.

In the presence of a magnetic field, passing a current through the device causes the direction of magnetization of a ferromagnetic film of nickel–iron alloy to rotate periodically relative to the magnetization of a cobalt–iron alloy separated by a thin layer of copper. This results in a periodic change in the voltage across the layered structure at microwave frequencies.

The device is much smaller than the cumbersome quartz crystals currently used as clocks or frequency standards in information processing and telecommunications, but has comparable stability. It could conceivably be fabricated at the nanoscale directly onto chips.

Philip Ball

Structural biology

Parasite protein probed

Structure 12, 41–53 (2004)

Chagas disease, an incurable and sometimes fatal condition, affects an estimated 18 million people in the Americas. As the disease progresses, the heart, oesophagus, colon and peripheral nervous system suffer irreparable damage, and death can occur suddenly.

The disease is caused by a parasite,

Trypanosoma cruzi, which passes to humans through infected blood transfusions or the faeces of some blood-sucking insects. Maria Harkiolaki *et al.* now unveil the X-ray structure of a protein that is essential to the parasite's survival.

The protein, dUTPase, is found in most organisms and is an enzyme involved in ensuring the integrity of DNA. Without it, cells die. Harkiolaki *et al.* show that the structure of *T. cruzi* dUTPase has a novel protein fold, making it very different from the human form. The finding may help in designing drugs that kill the parasite by turning the *T. cruzi* enzyme off, while leaving the human enzyme unharmed.

Helen R. Pilcher



Animal behaviour

Dead reckoning in the dark

Proc. Natl Acad. Sci. USA 101, 1105–1109 (2004)

Tali Kimchi and colleagues show that the blind mole rat (*Spalax ehrenbergi*) uses the Earth's magnetic field to guide its subterranean excursions. The researchers found that when the animals were placed in an artificially skewed magnetic field, the mole rats quickly lost their bearings.

Like all animals, blind mole rats need to navigate efficiently. In the absence of visual landmarks, they require a reliable compass to verify that they're moving in the direction they think they are (a process called path integration, or dead reckoning).

Kimchi *et al.* tested the rodents' navigational ability by presenting them with a wheel-shaped maze that had eight spokes, one of which led straight to the home nest. Mole rats left at the maze's hub, after a long walk around its perimeter, made straight for the nest. But when an artificial magnetic field was applied at 90° to the Earth's field, the animals made off in a correspondingly different direction.

The mole rats also used the Earth's magnetic field to find short-cuts in a rectangular grid, much like the streets in a US city. When faced with a zig-zag route to a food source, they subsequently devised a much shorter path, with less than half as many twists and turns. But when the magnetic field was rotated, the mole rats were less sure of themselves and frequently doubled back.

Michael Hopkin

TALI KIMCHI

Scent-triggered navigation in honeybees

Bees react to a perfume reminiscent of a distant food source by revisiting the site.

The honeybee, *Apis mellifera*, navigates rapidly and accurately to food sources that are often kilometres away^{1,2}. They achieve this by learning visual cues, such as the location and colour of nectar-bearing flowers¹⁻³, and chemical cues, such as the scent and the taste of the nectar^{1,2}. Here we train bees to visit differently scented sugar feeders placed at specific outdoor locations and find that they can be induced to visit the same locations simply by having the corresponding scent blown into the hive, even when the destinations no longer have the food or carry the scent. A familiar nectar scent can trigger specific memories of a route and therefore expedite navigation to the food source.

Von Frisch¹ observed that a floral scent on a forager bee returning to the hive from a natural food source can induce other experienced foragers to revisit that source without communicating with each other by waggle dancing⁴. This and similar observations^{2,5-7} raise the question of what memories are triggered by scent. We now investigate this in an experiment designed to ensure that the bees do not just home in on a familiar scent^{1,2,5-7} and are not recruited by waggle dances.

We trained individually marked honeybees to visit two sugar feeders positioned at two different locations, each 50 m from the hive; the distance between the two locations was 30 m. Each feeder carried a different scent, and feeders were offered alternately, switching every half hour. In the first experiment, scent one (offered at location one) was rose, and scent two (offered at location two) was lemon. After two days of training, the bees were tested periodically.

In the tests, the feeders were replaced by two empty, unscented jars. Scents one and two were blown into the hive, in turn, for 8 min each, by using a small fan attached to the hive entrance. During each scent-blowing interval, we noted the number of marked bees arriving at each test feeder. We also noted which feeder was visited first, the number of circlings (flights around the feeder), the number of landings on the feeder and the total number of visits (sum of all circlings and landings, including first visits). Tests were performed four times for each scent.

The trained bees that emerged from the hive upon injection of a scent showed a significant preference for the location that had carried the scent during training (Fig. 1). Injection of rose scent into the hive caused the trained bees to fly predominantly to location one. When lemon scent was blown into the hive, the overwhelming majority of trained bees visited location two. We repeated the

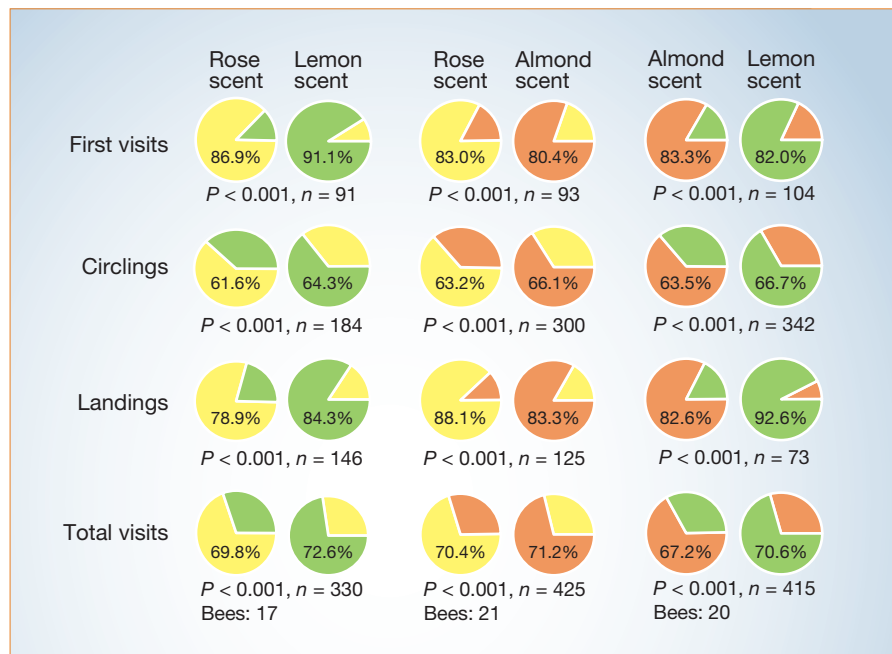


Figure 1 Scent association and foraging by honeybees. Three pairs of scent combinations were tested: rose and lemon, rose and almond, and almond and lemon. The proportion of first visits, circlings, landings and total visits to the two test feeders are shown. Colours indicate preferred feeder location: yellow, rose; green, lemon; orange, almond. The number of individually marked bees that visited both test feeders, the number of choices (n) and the P -values (χ^2 -test for choice frequencies) are shown at the bottom.

experiment with different pairs of scents: rose and almond (experiment two), and lemon and almond (experiment three). Each time we obtained similar results (Fig. 1). These findings indicate that each injected scent triggers visual memories of a specific location that the trained bees had previously visited.

Our test feeders were unscented, so the bees could not have found the correct site simply by homing in on the familiar scent, which is a well-known foraging strategy in bees^{1,2,8}. As the test feeders were empty, the arriving bees were not rewarded and so would not have danced upon return to the hive to recruit nestmates^{1,4}. Bees arriving at the correct location must therefore have relied on previously acquired navigational memories evoked by the injected scent.

Our findings reveal that associative recall is a mechanism used by honeybees, in addition to the dance language^{1,4}, to navigate successfully and repeatedly to an attractive food source. The taste and scent of nectar samples distributed by foragers on return to the hive¹ could trigger recall of specific visual and route memories associated with the food site in experienced recruits. Recalling factors such as the distance and direction of the food source, the landmarks en route and at the destination, and the flower's colour and shape would speed up flight to the food site, enhancing the foraging efficiency of the colony.

Judith Reinhard, Mandyam V. Srinivasan, Shaowu Zhang

Visual Sciences, Research School of Biological Sciences, The Australian National University, GPO Box 475, Canberra, ACT 2601, Australia
e-mail: reinhard@rsbs.anu.edu.au

1. von Frisch, K. *The Dance Language and Orientation of Bees* (Harvard Univ. Press, London, 1993).
2. Gould, J. L. in *Insect Learning* (eds Papaj, D. R. & Lewis, A. C.) 18–50 (Chapman & Hall, New York, 1993).
3. Wehner, R. in *Handbook of Sensory Physiology* (ed. Autrum, H.) 288–616 (Springer, Berlin, 1981).
4. Dyer, F. C. *Annu. Rev. Entomol.* **47**, 917–949 (2002).
5. Ribbands, C. R. *Proc. R. Entomol. Soc. A* **29**, 141–144 (1954).
6. Free, J. B. *Nature* **222**, 778 (1969).
7. Jakobsen, H. B., Kristjansson, K., Rohde, B., Terkildsen, M. & Olsen, C. E. *J. Chem. Ecol.* **21**, 1635–1648 (1995).
8. Tautz, J. & Sandeman, D. C. *J. Comp. Physiol. A* **189**, 293–300 (2003).

Competing financial interests: declared none.

Lifespan

Catch-up growth and obesity in male mice

Poor fetal growth is linked with long-term detrimental effects on health in adulthood¹. Here we investigate whether the lifespan of male mice is affected by their growth rate when they were suckling and find that limiting growth during that period not only increases longevity but also protects against the life-shortening effect of an

Table 1 Dietary factors and lifespan of male mice

Group	Pregnancy diet (% protein)	Lactation diet (% protein)	Weaning diet	Average age at death (days)
Normal chow	20	20	Chow	765 ± 22
Normal cafeteria	20	20	Cafeteria	715 ± 21
Catch-up chow	8	20	Chow	568 ± 36
Catch-up cafeteria	8	20	Cafeteria	517 ± 35
Postnatal low-protein chow	20	8	Chow	814 ± 25
Postnatal low-protein cafeteria	20	8	Cafeteria	807 ± 28

The different dietary regimes are summarized in the first three columns ($n = 24$ mice per group). Lifespans are expressed as mean ± standard error and were analysed by two-way analysis of variance followed by Duncan's post-hoc testing where appropriate. Effect of early diet: $P < 0.001$; effect of obesity, $P < 0.01$.

obesity-inducing diet later on. By contrast, we find that lifespan is considerably shortened if the postnatal period of growth is accelerated to make up for reduced growth *in utero*, and that, in addition, these mice are susceptible to the adverse effects on longevity of an obesity-inducing diet after weaning.

We compared the longevity of male mice that had received poor nutrition during either their fetal or early postnatal life, and determined the additional impact of an obesity-inducing diet. We fed pregnant mice either a normal (containing 20% protein) or a low-protein (8%) diet to restrict fetal growth. At birth, we cross-fostered pups so that the offspring of mothers fed on a low-protein diet during pregnancy (litters culled to four pups to maximize nutrition and rate of catch-up growth during lactation) were suckled by normally fed dams ('catch-up' animals)²; the pups of mothers fed normally during pregnancy (unculled to minimize nutrition and growth during lactation) were suckled by dams fed on the low-protein diet ('postnatal low protein' animals)². 'Normal' animals were the offspring of mothers fed on the 20%-protein diet during both pregnancy and lactation.

At 21 days of age, half of the pups from each litter were weaned onto standard laboratory chow and the others onto an obesity-inducing, 'cafeteria-style' diet³ (Table 1). This fattening diet contained 330 g standard laboratory chow, 330 g kg⁻¹ full-fat sweetened condensed milk, 70 g kg⁻¹ sucrose and 270 g kg⁻¹ water. The diet is used as a model for human obesity that is based on excess energy consumption, because its high palatability reliably leads to obesity.

Mice that underwent fetal-growth restriction but which were cross-fostered at birth to normally fed dams experienced rapid catch-up growth and died at a younger age than controls ($P < 0.001$; Table 1). In contrast, mice that grew normally *in utero* but were nursed by low-protein-fed dams showed increased longevity ($P < 0.01$; Table 1). Subsequent consumption of the cafeteria-style, obesity-inducing diet reduced life expectancy in control mice by 7% ($P < 0.01$) and in catch-up mice by 9% ($P < 0.05$). However, the cafeteria-style diet had no effect on the longevity of mice whose growth was restricted during suckling (Table 1). The greatest

difference in longevity was a 57% increase in lifespan ($P < 0.001$), which occurred between chow-fed, postnatally growth-restricted mice and catch-up mice fed on an obesity-inducing diet.

The mechanisms that underpin this large change in longevity are unclear. Reduced cell signalling by insulin-like peptides seems to increase the lifespan of nematodes, flies and mice⁴, and is thought to protect against stresses such as oxidative damage⁴. Accelerated shortening of chromosomal telomeres, due to increased cell division and/or oxidative damage, may occur in the catch-up group of mice (as found in rats²) and could lead to cell senescence in critical organs such as the kidneys^{2,5,6}.

We have shown that minor manipulation of maternal diet can increase life expectancy in mice by more than 50%, a discovery that calls for attention to this aspect of growth and nutrition in humans. There is, after all, a significant difference between living to be 50 years old and reaching the age of 75.

Susan E. Ozanne, C. Nicholas Hales

Department of Clinical Biochemistry, University of Cambridge, Cambridge CB2 2QR, UK
e-mail: seo10@cam.ac.uk

1. Hales, C. N. & Barker, D. J. P. *Br. Med. Bull.* **60**, 5–20 (2001).
2. Jennings, B. J., Ozanne, S. E., Dorling, M. W. & Hales, C. N. *FEBS Lett.* **448**, 4–8 (1999).
3. Petry, C. J., Ozanne, S. E., Wang, C. L. & Hales, C. N. *Clin. Sci.* **93**, 147–152 (1997).
4. Tatar, M., Bartke, A. & Antebi, A. *Science* **299**, 1346–1351 (2003).
5. Johnson, P. R., Stern, J. S., Horwitz, B. A., Harris, R. E. & Greene, S. F. *Am. J. Clin. Nutr.* **66**, 890–903 (1997).
6. Wright, W. E. & Shay, J. W. *Nature Biotechnol.* **20**, 682–688 (2002).

Competing financial interests: declared none.

Palaeogeography

Devonian tetrapod from western Europe

Several discoveries of Late Devonian tetrapods (limbed vertebrates) have been made during the past two decades^{1–3}, but each has been confined to one locality. Here we describe a tetrapod jaw of about 365 million years (Myr) old from the Famennian of Belgium, which is the first from western continental Europe. The jaw closely resembles that of *Ichthyostega*, a Famennian tetrapod hitherto known only

from Greenland. The environment of this fossil provides information about the conditions that prevailed just before the virtual disappearance of tetrapods from the fossil record for 20 Myr.

The tetrapod jaw described here was found in the nineteenth century at Strud, Belgium, in the Evieux Formation, and was initially presumed to be from a fish⁴. It shows typical tetrapod characters in its bone ornamentation and chamfered ventral margin of the dentary⁵ (Fig. 1a). The teeth are widely spaced and strongly curved posteriorly (Fig. 1b), as in the corresponding jaw portion of *Ichthyostega* from Greenland⁶. Although the find cannot be referred with certainty to *Ichthyostega*, it is strong evidence for the existence of a close relative of this tetrapod genus outside Greenland.

The Evieux Formation (Upper Famennian; *expansa* conodont zone) consists of shale, arkosic and evaporitic dolomite beds accumulated within 5–8-metre-thick sequences. The environment was fluvial to estuarine, and the uppermost part of the formation grades into a more marine environment with tempestite^{7,8}. The sandstone surrounding the jaw is indicative of fluvial conditions, as evidenced by a lack of grain sorting and by palaeosol fragments. This tetrapod therefore lived in rivers and estuaries, but the shoreline at the time of the Evieux Formation was oscillating south to north of Strud. The environment of the Devonian tetrapods seems to vary^{2,3,9,10}. For example, one genus was recovered from marine sequences², whereas the Greenland tetrapods *Ichthyostega* and *Acanthostega*, found in massive red sandstone, lived in rivers⁹, far to the north of the Devonian shore.

The biogeographical distribution of Devonian tetrapods is near-global for the group³ but extremely restricted (to a single locality or small geographic area) at the generic level, which was presumably a result of the colonization of continental environments that were less favourable to wide dispersal than were marine ones. The occurrence of an *Ichthyostega*-like form in Belgium extends the geographical distribution for at least one taxon and suggests that they were not as restricted as previously thought. The Famennian river system of the Ardennes, Belgium, is assumed to originate from western Germany⁸ and, although both Belgium and Greenland were part of the Euramerican continent, they were separated by at least 1,500 km.

The Late Devonian saw an increasing faunal exchange between Euramerica and Gondwana, leading to a loss of faunal regionalism. The rapid spread of tetrapods from a probable Euramerican cradle⁶ forms part of this picture, as does the arrival in Euramerica of fish groups of Gondwanan origin (notably the gyracanthid acanthodians and the lobefinned rhizodonts and megalichthyids).

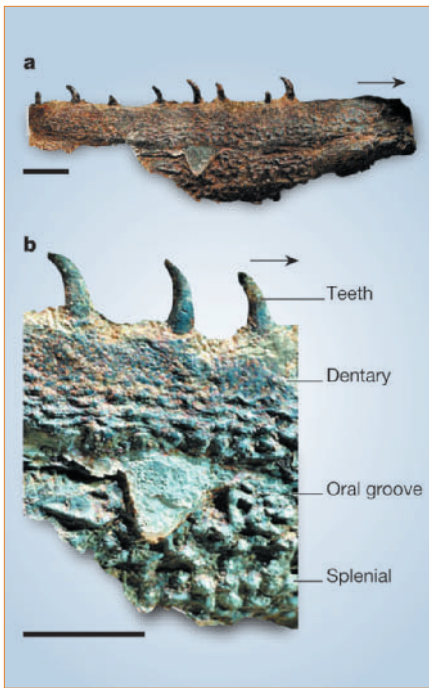


Figure 1 Lower jaw of a Late Devonian (Famennian) ichthyostegid tetrapod from Belgium (Liège University palaeontology collection, 6106a). **a**, External surface; **b**, detailed view of the teeth and ornamentation. Arrows point to anterior. Scale bar, 10 mm.

Until now, the record of these latter three groups in Euramerica has been restricted to Pennsylvania, near the southwestern extremity of the supercontinent¹¹, but the Strud ichthyostegid is associated with a rhizodont and a megalichthyid.

Western Europe, the central south coast of Euramerica, therefore seems to have had biogeographical links both with the northern inland (Greenland) and the southwest (Pennsylvania) in the latest Devonian. At the beginning of the Carboniferous, tetrapods virtually disappear from the fossil record for a 20-Myr interval known as Romer's gap⁹. The beginning of this gap coincides with the spread of rhizodonts, megalichthyids and gyracanthids across the whole of Euramerica, and with the extinction of the placoderm-dominated 'Old Red Sandstone fauna' associated with Devonian tetrapods in Greenland and elsewhere.

Although the causes of this faunal turnover remain uncertain, biotic factors associated with faunal exchange are a clear possibility. The Strud fossils give a tantalizing glimpse of Euramerican vertebrate biogeography immediately before this poorly understood phase of tetrapod evolution.

Gaël Clément*, Per E. Ahlberg†, Alain Blicke‡, Henning Blom§, Jennifer A. Clack§, Edouard Poty||, Jacques Thorez||, Philippe Janvier*¶

*UMR 5143, CNRS, Muséum National d'Histoire Naturelle, 75005 Paris, France, and †The Natural History Museum, London SW7 5BD, UK e-mail: janvier@mnhn.fr

‡Department of Evolutionary Organismal Biology,

Uppsala University, 752 36 Uppsala, Sweden
‡UMR 8014, CNRS, USTL Sciences de la Terre, 59655 Villeneuve d'Ascq Cedex, France
§University Museum of Zoology, Cambridge, Downing Street, Cambridge CB2 3EJ, UK
||Département de Géologie, Université de Liège, Sart-Tilman, 4000 Liège-1, Belgium

1. Ahlberg, P. E. & Milner, A. R. *Nature* **368**, 507–514 (1994).
2. Clack, J. A. *Gaining Ground* (Indiana University Press, Bloomington, 2002).
3. Zhu, M., Ahlberg, P. E., Zhao, W. & Jia, L. *Nature* **420**, 760–761 (2002).
4. Lohst, M. *Ann. Soc. Géol. Belg.* **15**, 112–203 (1888).
5. Ahlberg, P. E. & Clack, J. A. *Trans. R. Soc. Edinb. Earth Sci.* **89**, 11–46 (1998).
6. Jarvik, E. *Fossils Strata* **40**, 1–213 (1996).
7. Thorez, J., Steel, M., Bouckaert, J. & Bless, M. J. M. *Mededel. R. Geol. Dienst* **28**, 17–28 (1977).
8. Thorez, J., Goemaere, E. & Dreesen, R. in *Tide-influenced Sedimentary Environments and Facies* (eds De Boer, P. L. et al.) 389–415 (Reidel, Dordrecht, 1988).
9. Coates, M. I. & Clack, J. A. *Bull. Mus. Natl. Hist. Nat., Paris* **17**, 373–388 (1995).
10. Lukševičs, E. & Zupins, I. *Ichthyolith Issues, Spec. Publ.* **7**, 1302–1314 (2003).
11. Daeschler, E. B. & Shubin, N. *Nature* **391**, 133 (1998).

Competing financial interests: declared none.

COMMUNICATIONS ARISING

TNF ligands

Is TALL-1 a trimer or a virus-like cluster?

Native TALL-1 (B-cell activation factor, BAFF; also known as BlyS) was initially described as a homotrimer^{1–3}, but Liu and colleagues claim that it is a 60-subunit complex^{4–6} on the basis of their results from X-ray crystallography and size-exclusion chromatography. They consider TALL-1 60-mers to be the biologically active form,

and the arrangement of the 60-mers resembles that of the capsid of satellite tobacco necrosis virus^{4–6}. Here we show that active TALL-1 is trimeric under normal physiological conditions and that formation of higher-order oligomers is an artefact of tagging the amino terminus of the protein with a histidine tag.

A comparison of different TALL-1 constructs reveals that the choice of affinity tag influences the oligomerization state of the resulting protein. Proteins are trimeric when untagged^{1,3} or Myc-tagged², and form 60-mers when TALL-1 is histidine (His)-tagged^{4,5}; untagged TALL-1 purified as a trimer and bound to three TALL-1-receptor subunits crystallizes as a 60-mer⁶.

As fusion tags affect the oligomerization state of proteins⁷, we used size-exclusion chromatography to determine the oligomeric state of three different versions of a domain in TALL-1 that is homologous to tumour-necrosis factor (TNF-homology domain, or THD; residues 134–285): one untagged and two amino-terminally tagged THDs (Flag-TALL-1 and His-TALL-1). The latter two constructs link the tag and THD through a site that is cleaved by the protease factor Xa. Proteins were expressed in *Escherichia coli*, purified by affinity chromatography and assayed at pH 8.0 (Fig. 1a). According to Liu *et al.*⁴, TALL-1 is a trimer at pH 6.0 and a 60-mer at pH 7.4, so pH 8.0 should favour the formation of higher-order oligomers.

Our results show that the His-tagged protein forms an oligomer with an apparent relative molecular mass (M_r) of more than 900,000 (900K) (M_r of the His-TALL-1 60-mer would be 1,180K), confirming the

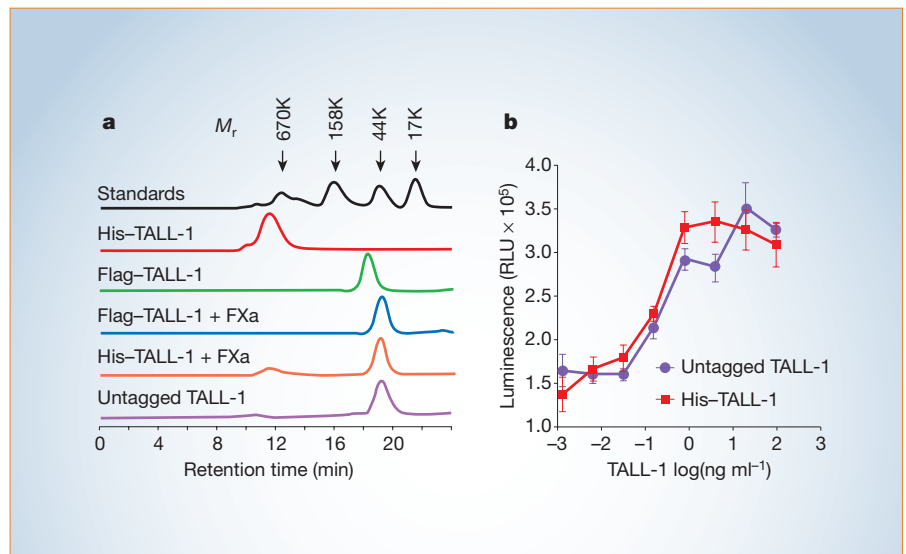


Figure 1 Effect of different amino-terminal tags on the oligomerization and biological properties of TALL-1. **a**, Size-exclusion chromatography (in 25 mM Tris, pH 8.0, 25 mM MgCl₂, 100 mM NaCl, 10% glycerol) using gel-filtration M_r standards (left to right: thyroglobulin, γ -globulin, ovalbumin and myoglobin). His-TALL-1 elutes as a higher oligomer (M_r > 900K). Flag-TALL-1 elutes as a trimer (62K). Flag-TALL-1 after proteolysis with factor Xa (FXa) yields untagged TALL-1 (44K). His-TALL-1 treated with FXa yields untagged protein (45K) and a small oligomeric fraction with an incompletely removed His tag. 'Untagged TALL-1' (45K) is from Biocarta. **b**, Viability of purified human B cells in the presence of untagged TALL-1 (purple) or His-TALL-1 (red), using anti-IgM antibody (2.0 mg ml⁻¹) as co-stimulator. RLU, relative luminescent units.

results of Liu *et al.*^{4,5}. However, Flag-TALL-1 elutes as a trimer of 62K (M_r of the Flag-TALL-1 trimer is 58K) that is stable for several months. Liu *et al.*⁴ used a western blot of fractions from size-exclusion chromatography to resolve both trimeric and oligomeric forms of Flag-TALL-1. However, the latter species eluted in the void volume of a Superdex-200 gel-filtration column, which is consistent with either the 60-mer or non-specific aggregates.

We also tested whether the Flag tag prevents the formation of 60-meric TALL-1 and whether the His tag causes trimeric TALL-1 to form 60-mers. We removed the TALL-1 tags using factor Xa and analysed the size of the digested products. We found that removing the Flag tag did not alter TALL-1's oligomerization state. Removing the His tag caused TALL-1 to convert from a 60-mer to a trimer (Fig. 1a). This suggests that the His tag promotes formation of non-native oligomers.

The pH dependence of the trimer-to-60-mer transition⁴ alters with the ionization of the imidazole ring of histidine (pKa 6.0–7.0). Whereas a neutral amino-terminal tag (that is, when the histidine ring is not ionized because the buffer pH is above its pKa) promotes 60-mer formation, a positively charged His tag (buffer pH below the histidine pKa) or a negatively charged Flag tag does not induce TALL-1 oligomerization. Consistent with this, untagged TALL-1 is trimeric (Fig. 1a).

We tested viability (Fig. 1b) and B-cell proliferation (data not shown) and confirmed that trimeric TALL-1 is biologically active. It is not surprising that trimers and 60-mers are both active because oligomerization does not always affect biological activity⁸; alternatively, interaction with membrane-bound TALL-1 receptors may force dissociation of 60-mers, such that both species give a similar biological response.

Our results indicate that native TALL-1, in common with all members of the TNF-ligand superfamily, is trimeric and biologically active. We have shown that introduction of an amino-terminal histidine tag affects the oligomerization state of TALL-1, and we warn generally against the use of fusion tags that might perturb the physical chemical properties of the host protein.

Eugene A. Zhukovsky*, Jie-Oh Lee†, Michael Villegas*, Cheryl Chan*, Seung Chu*, Cameron Mroske*

*Xencor, Monrovia, California 91016, USA

e-mail: ezhukovsky@xencor.com

†Department of Chemistry and School of Molecular Science, KAIST, Daejeon 305–701, Korea

1. Kanakaraj, P. *et al.* *Cytokine* **12**, 25–31 (2001).
2. Karpusas, M. *et al.* *J. Mol. Biol.* **315**, 1145–1154 (2002).
3. Oren, D. A. *et al.* *Nature Struct. Biol.* **9**, 288–292 (2002).
4. Liu, Y. *et al.* *Cell* **108**, 383–394 (2002).
5. Liu, Y. *et al.* *Nature* **423**, 49–56 (2003).
6. Kim, H. M. *et al.* *Nature Struct. Biol.* **10**, 342–348 (2003).
7. Wu, J. & Filutowicz, M. *Acta Biochim. Pol.* **46**, 591–599 (1999).
8. Arunkumar, A. *et al.* *Protein Sci.* **11**, 1050–1061 (2002).

Liu *et al.* reply — Our conclusion that a 60-meric soluble-TALL-1 cluster is the functional unit of this ligand^{1,2} is based on several lines of evidence.

We have shown that soluble TALL-1 (sTALL-1; residues 134–285) with a Flag tag, secreted from 293 cell lines, spontaneously forms a virus-like cluster in the culture medium under physiological conditions¹. Also, Flag-tagged proteins dissociate into trimers and monomers at low pH, ruling out the possibility of random aggregation that is suggested by Zhukovsky and colleagues.

The pH dependence of the sTALL-1 structure is consistent with the presence of a histidine residue in the flap region¹, the removal of which prevents the oligomerization of sTALL-1 and abolishes the activity of the ligand in transfection and in B-cell-stimulation assays¹. However, this truncated version of the molecule still binds to its receptor¹, indicating that the missing flap region is not involved in receptor binding². The structure of the truncated molecule is the same as that of native sTALL-1, apart from the missing flap region².

It is true that some His6-tagged proteins are susceptible to aggregation caused by divalent metals. We tested whether this could be the case for our His6-tagged sTALL-1 protein by adding 100 mM EDTA, pH 7.5, at the different stages of purification. We found that it did not affect oligomerization. Note also that the His6 tag is far away from the trimer–trimer association region in our structure¹.

To prevent any perturbation of the oligomerization properties of sTALL-1 by the His6 or Flag tags, we overexpressed an untagged version of sTALL-1 (residues 134–285) in *Escherichia coli*. This sTALL-1 (prepared and characterized with the help of Zhongzhou Chen) was purified on a Q-Sepharose column in buffer (20 mM Tris-HCl) at pH 7.2, concentrated in an Amicon Ultra unit (Millipore) and applied to Superdex-200 (150 mM NaCl, 20 mM Tris-HCl) at pH 8.0. sTALL-1 emerged as a multimer in the void volume (data not shown). The buffer containing this sample was then adjusted to pH 6.0, equilibrated for 2 hours and reappplied to the Superdex-200 column in the same buffer at pH 6.0.

The elution profile consisted of a single sharp peak: the molecular size of the material in this fraction was estimated from the elution positions of three M_r standards (blue dextran, 2,000,000 (2,000K); ferritin, 440K; and ovalbumin, 43K) under the same conditions. The pure sTALL-1 protein eluted at a position corresponding to a trimer. Amino-terminal sequencing analysis confirmed that this was the correct protein (data not shown).

Other crystal structures^{3,4} reported for trimeric sTALL-1 were obtained with pro-

teins that had been crystallized at low pH (pH 4.5 and 6.0, respectively), conditions under which the trimer should be the only species present¹. Zhukovsky *et al.* used native sTALL-1, with and without tags, in their assays. As these should both be transformed to 60-mers under physiological conditions, it is not surprising that these forms are active. If, by contrast, they had used a flapless version of the protein, which does not oligomerize, we predict that it would have been inactive. We contest that neither the His6 tag nor the Flag tag promotes or blocks the formation of 60-mers by sTALL-1 and that it is the 60-mer, rather than a trimer or some other oligomer, that is the functional unit.

Other data indicate that further oligomerization of ligand trimers (multiple valences) or multiple trimers are required to activate the signal-transduction pathway of the TNF superfamily. For example, TNF receptors can oligomerize on the cell surface without bound ligand⁵. TNF-receptor-associated factors (TRAF), which mediate downstream signal transduction by the TNF family, exist as trimers before TNF ligands are recruited to its receptors⁶. Ectodysplasin ligand A1 (EDA-A1) and A2 (EDA-A2) require a collagen-like region for their proper function, which suggests that two dimers or higher-order clusters are needed for the activation of these ligands⁷. Combining with our results with sTALL-1, it is reasonable to suggest that one trimer of the ligand is not enough to trigger signal transduction for TNF family members. Multiple trimers (including trimers in different oligomerization states of increased concentration and valence and hence increased avidity) or a local accumulation of trimers (increasing in concentration alone) on the cell surface may be required for the optimal recruitment of a cluster of receptors to activate signal transduction. This is similar to the supermolecular activation cluster phenomenon of T-cell activation that occurs after binding to antigen-presenting cells⁸.

Xia Hong*, John Kappler*†, Yingfang Liu*, Liangguo Xu*, Hong-Bing Shu*, Gongyi Zhang*

*Integrated Department of Immunology, National Jewish Medical and Research Center, †Howard Hughes Medical Institute, and Department of Pharmacology, Biomolecular Structure Program, School of Medicine, University of Colorado Health Science Center, Denver, Colorado 80206, USA
e-mail: zhangg@njc.org

1. Liu, Y. *et al.* *Cell* **108**, 383–394 (2002).
2. Liu, Y. *et al.* *Nature* **423**, 49–56 (2003).
3. Karpusas, M. *et al.* *J. Mol. Biol.* **315**, 1145–1154 (2002).
4. Oren, D. A. *et al.* *Nature Struct. Biol.* **9**, 288–292 (2002).
5. Chan, F. K. *et al.* *Science* **288**, 2351–2354 (2000).
6. Park, Y. C. *et al.* *Nature* **398**, 533–538 (1999).
7. Srivastava, A. K. *et al.* *Proc. Natl Acad. Sci. USA* **94**, 13069–13074 (1997).
8. Monks, C. R., Freiburg, B. A., Kupfer, H., Siaty, N. & Kupfer, A. *Nature* **395**, 82–86 (1998).

Summing up the noise in gene networks

Johan Paulsson

Department of Molecular Biology, Princeton University, Washington Road, Princeton, New Jersey 08544-1014, USA, and Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, Wilberforce Road, University of Cambridge, Cambridge CB3 0WA, UK

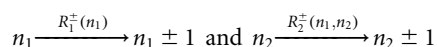
Random fluctuations in genetic networks are inevitable as chemical reactions are probabilistic and many genes, RNAs and proteins are present in low numbers per cell. Such 'noise' affects all life processes and has recently been measured using green fluorescent protein (GFP). Two studies show that negative feedback suppresses noise, and three others identify the sources of noise in gene expression. Here I critically analyse these studies and present a simple equation that unifies and extends both the mathematical and biological perspectives.

Intracellular randomness has long been predicted from basic physical principles¹ and observations of phenotypic heterogeneity^{2,3}. In the last few years it has also been visualized directly using fluorescent probes. The first quantitative studies^{4–8} collectively examined the noise associated with the principal steps of the central dogma of molecular biology; that is, replication, gene activation, transcription, translation and the enslaving intracellular environment. They also suggested how autorepression of replication and transcription suppresses noise, and how eukaryotes differ from prokaryotes. This analysis connects the different studies to a simple variant of the fluctuation-dissipation theorem and uses the experimental controls to extend or reinterpret many of the conclusions.

Theory

The fluctuation-dissipation theorem

All cell components display intrinsic noise due to random births and deaths of individual molecules, and extrinsic noise due to fluctuations in reaction rates. This has been modelled extensively using both detailed computer simulations and analytically tractable Markov processes, each tailored to a particular system. Here, I take a different approach and see how far a first approximation of a simple but generic model goes in explaining previous theory and experiments. The probabilities of having n_1 and n_2 molecules per cell of chemical species X_1 and X_2 (for instance messenger RNAs and proteins, but interpretations vary with application) are described by a birth-and-death Markov process with events



Because n_1 affects rate R_2 but n_2 does not affect R_1 , this is an example of dynamic disorder^{9,10} where species X_1 provides the randomly fluctuating environment for X_2 , as mRNA fluctuations randomize protein synthesis. To collectively approximate all such processes I use the Ω -expansion^{11,12} where the first- and second-order terms reproduce the macroscopic rate equations and the fluctuation-dissipation theorem^{13,14} respectively. The latter is then interpreted in terms of the logarithmic gains $H_{ij} = \partial \ln(R_i^-/R_i^+)/\partial \ln(n_j)$ that measure how the balance between production and elimination of X_i is affected by X_j . These scale-free parameters are closely related to the elasticities of metabolic control analysis^{15,16} and the apparent kinetic orders of biochemical systems theory^{17,18}, and can often be estimated directly from the reaction rates. For instance, if R_i^+ and R_i^- are of first and second kinetic order in n_j respectively, then $H_{ij} = 2 - 1 = 1$. For the process described above, using σ_i for standard deviations, $\langle n_i \rangle$ for averages and τ_i for average lifetimes, stationary fluctuations around a stable fixed point

follow

$$\frac{\sigma_2^2}{\langle n_2 \rangle^2} \approx \underbrace{\frac{1}{\langle n_2 \rangle H_{22}}}_{\text{Intrinsic } n_2 \text{ noise}} + \underbrace{\frac{\sigma_1^2}{\langle n_1 \rangle^2} \frac{H_{21}^2}{H_{22}^2}}_{\substack{n_1 \text{ noise} \\ \text{Susceptibility}}} \underbrace{\frac{H_{22}}{H_{11} / \tau_1 + H_{22} / \tau_2}}_{\text{Time-averaging}} \quad (1)$$

where $\sigma_1^2/\langle n_1 \rangle^2 \approx (\langle n_1 \rangle H_{11})^{-1}$. Intrinsic noise depends on the average number of molecules and how systematic adjustments (rate H_{22}/τ_2) quench spontaneous fluctuations (rate $1/\tau_2$). The normalized adjustment rate H_{22} can also be interpreted as the statistical bias to return to the average rather than deviate further: a 1% increase in n_2 gives a H_{22} per cent increase in R_2^-/R_2^+ . Extrinsic noise instead depends on the magnitude of n_1 fluctuations and how strongly n_1 affects n_2 . The normalized susceptibility factor H_{21}/H_{22} reflects that a 1% increase in n_1 gives a H_{21} per cent increase in R_2^-/R_2^+ , which makes n_2 adjust towards a H_{21}/H_{22} per cent lower average quasi-steady-state. When n_1 changes rapidly (high H_{11}/τ_1) or n_2 adjusts slowly (low H_{22}/τ_2), n_2 does not have time to reach its quasi-steady-state before n_1 changes anew. Consecutive ups and downs in n_1 then cancel out and n_2 time-averages over the recent history of n_1 fluctuations. The effect of cell growth and division is qualitatively accounted for by adding first-order elimination terms to R_1^- or R_2^- . The method behind equation (1) can be extended to any chemical system, providing a basis for a stochastic Biochemical Systems Theory (J.P., manuscript in preparation).

Noise in the central dogma

The basic principles of noise in genetic networks can be understood by applying equation (1) to the central dogma of molecular biology. Unregulated replication combined with first-order elimination is dynamically unstable as DNA acts as a template for its own synthesis: if each molecule on average replicates more or less than once per cell cycle, the average concentration increases until resources become limiting or decreases until all templates are gone. At intermediate replication frequencies, where synthesis and elimination are delicately balanced, random fluctuations still accumulate in an almost unrestrained fashion. With X_2 as self-replicators, this has been modelled¹⁹ similarly to equation (1) with $H_{22} \ll 1$, and the same principles apply to dynamic instability of microtubules²⁰ and many other cell processes^{12,19,21}. Unregulated transcription and translation are not unstable when combined with first-order elimination, as mRNAs and proteins are not templates for their own synthesis. Over the last 35 years, models of stochastic gene expression^{19,22–31} have instead focused on dynamically disordered linear processes, corresponding to $H_{11} = H_{22} = -H_{21} = 1$ in equation (1) with X_2 as proteins and X_1 as genes^{27,28}, mRNAs^{22–26} or the intracellular environment¹⁹. To

suppress the noise in any of these systems, cells commonly use autorepression that increases and decreases synthesis at low and high concentrations respectively. This has been studied extensively using macroscopic models^{17,32}, and the stochastic principles are closely related. Autorepression can raise the effective H_{22} , and thus: (1) increase the adjustment rate H_{22}/τ_2 relative to the rate $1/\tau_2$ of spontaneous randomization and thereby suppress intrinsic noise around a given average number of molecules; (2) increase the adjustment rate H_{22}/τ_2 relative to the rate H_{11}/τ_1 of environmental changes and thereby amplify extrinsic noise by preventing time-averaging; (3) decrease the susceptibility H_{21}/H_{22} of the quasi-steady-state, which typically overcompensates for the impaired time-averaging and produces a net decrease in extrinsic noise. This may explain the popularity of autorepression in transcription networks^{32–34} and its ubiquity in replication control of chromosomes and plasmids where it has been similarly described^{3,19}. Equation (1) thus unifies models of autoreplication, constitutive transcription and translation, and the stabilizing effect of autorepression, all in disordered environments. This makes it ideally suited also to unify the GFP studies that examine these aspects experimentally.

Terminology and measures

Comparing different systems requires consistency in definitions and measures. The terms ‘intrinsic’ and ‘extrinsic’ generically distinguish between the origin and propagation of noise, and their biological meaning is always defined in conjunction with a specified component or process. For example, if a gene for a transcriptional repressor spontaneously switches on and off, thereby enslaving the encoded protein and transmitting the fluctuations to repressed genes, the noise is intrinsic to the number of active repressor genes and extrinsic to all affected components. If the corresponding noise in a repressed protein was instead assigned to its transcription, just because transcription transmits the noise from the repressor gene, then by the same logic it must also be assigned to translation or to any other step in the cascade. This relates directly to the mathematical measures. For intrinsic noise it is convenient to use $\sigma_2^2/\langle n_2 \rangle$ for a size-independent comparison, but this artificially forces extrinsic noise to increase with $\langle n_2 \rangle$, as in $\sigma_2^2/\langle n_2 \rangle \approx H_{22}^{-1} + E/\langle n_2 \rangle$ where E is the second term in equation (1). Unless the measure matches the noise, scale artefacts may thus completely distort interpretations of dynamics. For instance, if all protein (X_2) noise came from fluctuations in protease levels, $\sigma_2^2/\langle n_2 \rangle^2$ would typically be independent of transcription and translation rates. But because both of these processes affect $\langle n_2 \rangle$, measuring noise strength by $\sigma_2^2/\langle n_2 \rangle$ would make it appear as if they contributed noise, even though they have nothing to do with either its production or transmission. Fortunately, these scaling principles also provide an opportunity to trace the noise: increasing the number of repressor genes in the example above should reduce the relative standard deviations of the repressed proteins, whereas other changes in transcription or translation should have little effect. This approach must be used carefully as the susceptibilities or time constants may change as well, but it at least provides useful indications.

Autorepression

Two groups have used plasmid-expressed *gfp* to examine, for the first times, how autorepression of replication⁴ and transcription⁵ affects noise levels. The replication study used GFP as a reporter for gene dosage, and compared natural and synthetic plasmids that have intact and impaired autorepression respectively. The transcription study instead used plasmids as cloning vectors to study protein fluctuations, and engineered an autorepression loop by placing a fusion of GFP and the tetracycline repressor (TetR) downstream of a TetR-repressed promoter. Both studies showed that autorepression substantially reduces relative standard deviations, which has been suggested to come from the more rapid adjustments to steady state in plasmid¹⁹ and protein⁵ concentrations respectively. But even if

rapid adjustments reduce intrinsic noise around a given average, the effect is the opposite for extrinsic noise (compare points (1) and (2) above). This reveals an interesting discrepancy between the studies: if the noise came from gene expression, GFP would not measure plasmid copy numbers, and if it came from fluctuations in plasmid copy numbers, a protein that adjusted more rapidly to its mean-dering steady state would inherit more noise, not less. Thanks to the many experimental controls, this issue can be at least partially settled using equation (1).

Replication

With X_2 as plasmids, the simplest models of plasmid copy number control¹³⁵ assume $R_2^+ \propto n_2(K + n_2^h)^{-1}$. Factor n_2 reflects that each copy can self-replicate, and factor $(K + n_2^h)^{-1}$ reflects autorepression, for instance approximating the effect of short-lived plasmid-encoded inhibitors¹⁹ that act cooperatively with effective Hill coefficient h . The statistics of plasmid segregation in growing cells depends on the accuracy of partitioning plasmids at cell division, but the qualitative effect is similar to first-order elimination¹⁹, $R_2^- \propto n_2$. When $K + n_2^h$ is insensitive to n_2 , so that autoreplication is poorly constrained by autorepression, both R_2^- and R_2^+ are thus approximately proportional to n_2 and plasmids are almost unstable, $H_{22} \ll 1$. This explains the large copy number variation observed for plasmids with impaired replication control⁴.

The random environment X_1 could represent DNA polymerase, ribosomes, chaperones or any other factor that affects the replication control system. But because equation (1) is generic, the qualitative effects of these hypothetical fluctuations can be understood even if the kinetic details are left unspecified. As in points (1) to (3) in the ‘Noise in the central dogma’ section, autorepression should suppress intrinsic plasmid noise by expediting normalized plasmid adjustments, and suppress extrinsic plasmid noise by decreasing the net influence of environmental variation. The six natural plasmids studied⁴ displayed similar $\sigma_2/\langle n_2 \rangle$ despite large differences in $\langle n_2 \rangle$ —typically the signature of extrinsic noise. However, low- and high-copy plasmids are different altogether, and the former are thought to compensate by more efficient partitioning and replication control. Some of the similarities also reflect that GFP is an inexact reporter of single-cell plasmid copy numbers (Supplementary Information). This makes it impossible to say how the noise is checked by autorepression, but because plasmids are present in low numbers per cell, most studies¹⁹ have focused on intrinsic noise and normalized adjustment rates.

Transcription

With X_2 instead representing proteins, constitutive transcription (R_2^+ independent of n_2) combined with first-order elimination ($R_2^- \propto n_2$) is far from unstable ($H_{22} = 1$). Because GFP was also present in high numbers per cell ($\langle n_2 \rangle \gg 1$), intrinsic protein noise should thus be negligible, as confirmed by experimental controls⁵ where $\sigma_2/\langle n_2 \rangle$ was unaffected by large changes in $\langle n_2 \rangle$. Similarly, experiments⁵ where $\sigma_2/\langle n_2 \rangle$ was unaffected by 20-fold changes in total gene activity and mRNA levels suggest that spontaneous small-number fluctuations in these components do not contribute any substantial protein noise, although they may still transmit fluctuations from plasmids, ribosomes or RNAses. The transcriptional autorepression loop was described⁵ by $R_2^+ \propto K(K + n_2)^{-1} \approx Kn_2^{-1}$, which would stabilize the system twofold ($H_{22} \approx 2$). The protein now affects the number of active genes and mRNAs, but because these components were not the source of the extrinsic noise, equation (1) still applies (R_1 is unaffected by n_2) and predicts points (2) to (3) in the ‘Noise in the central dogma’ section with the environment X_1 left unspecified. Rapid protein adjustments thus increase noise by preventing time-averaging, and the observed decrease instead probably comes from an overcompensating reduction in the susceptibility of the steady state to changes in some extrinsic factor. For instance, cells that by chance have more

plasmid copies accumulate more protein, but transcriptional autorepression then partially compensates by reducing the expression per gene copy. Experimental controls⁵ measuring the average protein level (n_2) as a function of the average plasmid level (n_1) suggest an expected $\langle n_2 \rangle \propto \langle n_1 \rangle$ and $\langle n_2 \rangle \propto \sqrt{\langle n_1 \rangle}$ for the unregulated and autoregulated system respectively, corresponding to $H_{21}^2/H_{22}^2 = 1$ and $H_{21}^2/H_{22}^2 \approx 1/4$ in equation (1). This does not prove that all protein noise comes from fluctuations in plasmid copy numbers—the results are also consistent with a decreased susceptibility to ribosomes or proteases—but plasmids are likely contributors because of their low copy numbers and because their fluctuations are slow and difficult to time-average. Many natural plasmids in fact use transcriptional autorepression³⁶ to ensure that their replication proteins are unsusceptible to perturbations in both plasmid copy numbers and the intracellular environment. The transcription study⁵ is the first direct assessment of this principle.

Chromosomal gene expression

Plasmid-expressed *gfp* does not reflect the randomness of chromosomal gene expression because transcription and translation are averaged over more gene copies and plasmids themselves fluctuate randomly. The first group⁶ to identify the sources of noise in gene expression therefore cloned *gfp* into the chromosome of *Bacillus subtilis* and measured single-cell fluorescence at varying rates of transcription and translation. The result of that study is consistent with a long-standing hypothesis^{19,22–26} that protein fluctuations depend on the burst b of proteins (henceforth denoted X_2) made per mRNA transcript, $\sigma_2^2/\langle n_2 \rangle \approx 1 + \langle b \rangle$. A second group⁷ instead inserted two types of *gfp* into the *Escherichia coli* chromosome and used correlations between them to infer where the fluctuations came from. Finally, a third group revisited the one-gene strategy for *Saccharomyces cerevisiae* and again reported $\sigma_2^2/\langle n_2 \rangle$ as a function of transcription and translation rates. They suggested that eukaryotes differ from prokaryotes because promoter fluctuations and transcriptional reinitiation produce a non-monotonous transcription noise.

Transcription and translation noise in *B. subtilis*

With X_1 and X_2 as mRNAs and proteins, and k_1 and k_2 as the transcription and translation rates, the model behind the *B. subtilis* experiments^{6,26} assumes $R_1^+ = k_1$, $R_1^- = n_1/\tau_1$, $R_2^+ = k_2 n_1$ and $R_2^- = n_2/\tau_2$, so that $H_{11} = H_{22} = -H_{21} = 1$ in equation (1). Further assuming $\tau_1 \ll \tau_2$ gives $\sigma_2^2/\langle n_2 \rangle^2 \approx \langle n_2 \rangle^{-1} + \langle n_1 \rangle^{-1} \tau_1/\tau_2$, and multiplying by $\langle n_2 \rangle = \langle n_1 \rangle k_2 \tau_2$ reproduces the burst prediction $\sigma_2^2/\langle n_2 \rangle \approx 1 + k_2 \tau_1 = 1 + \langle b \rangle$. The experiments confirmed that $\sigma_2^2/\langle n_2 \rangle$ is independent of k_1 and increases linearly with k_2 , which is quite remarkable as there are no free parameters and they tested a real prediction rather than making *ad hoc* explanations. This suggested that noise is determined translationally and that strongly translated genes are implicated in particularly noisy processes⁶. However, the burst term $\langle b \rangle$ in $\sigma_2^2/\langle n_2 \rangle$ does not come from the randomness of translation, but from random births and deaths of mRNA transcripts. It would disappear completely if transcription and mRNA degradation were deterministic, and would remain unchanged if translation was deterministic. By construction, $\sigma_2^2/\langle n_2 \rangle$ is thus independent of the contribution from random translation events, and using the scale-free $\sigma_2^2/\langle n_2 \rangle^2$ provides a quite different perspective^{19,31}: a higher translation rate has a marginal effect because it only reduces the intrinsic protein noise $\langle n_2 \rangle^{-1}$, which should be small throughout the experiments, whereas a higher transcription rate additionally reduces the much greater noise contribution from having few mRNA transcripts, $\sigma_1^2/\langle n_1 \rangle^2 = \langle n_1 \rangle^{-1}$.

The effect of other random influences was not estimated, and noise was measured at varying rates—tracking changes rather than absolute values—to draw conclusions independently of such effects⁶. But extrinsic processes would typically make superimposable contributions to $\sigma_2^2/\langle n_2 \rangle^2$ rather than to $\sigma_2^2/\langle n_2 \rangle$, and using $\sigma_2^2/\langle n_2 \rangle$ can then be misleading (see ‘Terminology and measures’ section). In

particular, if any substantial noise changed as transcription and translation were varied, total noise should behave in a more complicated manner and the experiments would be hard to interpret at all. If it instead remained constant, it could be calculated from the data. The acid test is how $\sigma_2^2/\langle n_2 \rangle^2$ appears to vary with $\langle n_2 \rangle^{-1}$ in the transcription experiments. The mRNA–protein model above predicts a perfect proportionality, a constant extrinsic noise adds a displacement, and a varying extrinsic noise creates a distortion. Replotted this way, the data fit well to a straight line with a small displacement, which suggests that most noise indeed comes from having few transcripts per cell, at least at low rates of transcription. This strengthens the authors’ model^{6,26}, although from a different biological perspective, and shows that their data additionally account for other possible sources of noise. Noise in promoter activity or fluctuations in the transcriptional repressor should also have less effect at higher transcriptional induction, and could in principle produce the same result; however, there is then no reason to expect a quantitative fit without fine-tuned parameters (Supplementary Information).

Intrinsic and extrinsic noise in *E. coli*

The kinetic analysis³¹ accompanying the two-gene *E. coli* experiments uses separate models for the noise that is intrinsic and extrinsic to the expression of an individual gene. The intrinsic part is similar to the mRNA–protein model above, and the extrinsic part is equivalent to the second term in equation (1) with arbitrary $\sigma_2^2/\langle n_2 \rangle^2$ and H_{21} , but $H_{11}/\tau_1 = 0$. The difference reflects that equation (1) is derived from stationary stochastic processes rather than fixed probability distributions, thereby extending the analysis from static ($H_{11}/\tau_1 = 0$) to dynamic ($H_{11}/\tau_1 > 0$) disorder. This is biologically relevant as GFP degrades so slowly that few cell processes are static in comparison.

Experimentally, Elowitz *et al.*⁷ measured the concurrent expression of two identically regulated but fluorescently distinguishable *gfp* genes in the *E. coli* chromosome. Spontaneous small-number mRNA fluctuations then generate uncorrelated fluctuations in the two GFPs as each allele makes its own transcript, whereas noise in RNA polymerase levels perturbs both GFPs in the same way. To quantify these data without relying on any particular kinetic model, Swain *et al.*³¹ described generic protein fluctuations in terms of distributions of underlying (static) variables that were either intrinsic (for instance mRNA) or extrinsic (for instance RNA polymerase) to the expression of an individual gene. This use of ‘intrinsic’ and ‘extrinsic’ makes a non-trivial distinction between system and environment, but is qualitatively supported by equation (1) where the second term can be interpreted as the normalized covariance $\sigma_{23}/(\langle n_2 \rangle \langle n_3 \rangle)$ between two identical and independent components X_2 and X_3 embedded in the same environment X_1 (Supplementary Information). Nonlinear mechanisms would violate the requirement that the GFPs are independent (Supplementary Information), but that was not a problem in the experiments, and this clever strategy then directly separated the noise that is intrinsic and extrinsic to the expression of an individual gene. Their intrinsic contribution behaved almost like the total noise in the one-gene *B. subtilis* study and was similarly explained in terms of small-number mRNA fluctuations. However, their extrinsic contribution from the intracellular environment was by contrast substantial or even dominant. This is a central discrepancy. If most noise comes from the low numbers of mRNAs or proteins, each gene is under individual selection for high or low noise production, but if it comes from the intracellular environment, numerous genes are affected similarly regardless of transcription and translation rates. Discrepancies can always be ascribed to differences in experimental conditions, but in this case the devil seems rather to be in the genetic details: most noise extrinsic to gene expression in the two-gene study was traced back^{7,31} to the transcriptional repressor LacI that was gradually inactivated by isopropyl- β -D-thiogalactoside to

vary the transcription rate. The two-gene study used wild-type, oscillating or plasmid-borne *lacI*, all of which are expected to be noisier than the highly expressed chromosomal *lacI* of the one-gene study. This more complex behaviour of the noise underscores the importance of using two GFPs, especially as $\sigma_2^2/\langle n_2 \rangle^2$ was a non-monotonic function of transcription.

Transcription and translation noise in *S. cerevisiae*

The detailed computer simulations accompanying the *S. cerevisiae* study⁸ account for numerous steps in eukaryotic gene expression, and the results perfectly match experiments where noise strength, as measured by $\sigma_2^2/\langle n_2 \rangle^2$, responded non-monotonically to transcriptional induction and increased linearly with translation rate. The translation response was much stronger at partially induced rather than maximally induced transcription. This suggested that, in addition to translation noise, eukaryote-specific chromatin remodelling and transcriptional reinitiation produce quantile transcription bursts. However, because $\sigma_2^2/\langle n_2 \rangle^2$ introduces a scale-dependence, the experiments should be re-evaluated in terms of $\sigma_2^2/\langle n_2 \rangle^2$. The published transcription data then show a monotonically decreasing $\sigma_2^2/\langle n_2 \rangle^2$, but the authors' full data set is non-monotonic and displays an initial increase in $\sigma_2^2/\langle n_2 \rangle^2$ at low induction (W. Blake, personal communication). The translation experiments basically reiterate this result: the linear increase in $\sigma_2^2/\langle n_2 \rangle^2$ indicates an extrinsic source of noise, and the stronger translation response was obtained at intermediate transcriptional induction where $\sigma_2^2/\langle n_2 \rangle^2$ was larger. This is still consistent with slow chromatin remodelling and transcriptional reinitiation, so their interpretation may be correct, but both monotonic and non-monotonic responses in either $\sigma_2^2/\langle n_2 \rangle^2$ or $\sigma_2^2/\langle n_2 \rangle^2$ are also perfectly consistent with a variety of other noise sources, such as repressors³¹ and activators or their random association and dissociation to DNA (Supplementary Information). Other types of experiments are thus needed to distinguish between the many alternatives, and at this point there are no reasons to favour one interpretation over another. So far, no principal differences between protein fluctuations in pro- and eukaryotes have been demonstrated. The claim that there is a difference was based on a comparison with the *B. subtilis* experiments⁶ that used the same methods and interpretations, whereas the *E. coli* experiments⁷ used other methods and interpretations but reported almost identical results. Instead, a comparison of the studies shows how dependent protein noise is on rate constants and upstream inputs; that is, differences in network design have overshadowed organismic differences. Notably, all of them indicate very little noise from, for example, RNA polymerase or ribosomes that would affect gene expression on a global level. In the one-gene experiments, fully induced genes displayed relative protein variation as low as 15% in both *B. subtilis* and *S. cerevisiae*, and much of this may reflect cell-size variation and measurement errors (Supplementary Information). In the two-gene *E. coli lacI⁻* experiments, the noise extrinsic to gene expression was so small (5%) that it cannot be convincingly distinguished from zero.

Outlook

Noise in genetic and metabolic networks can be detrimental to fitness—randomizing developmental pathways, disrupting cell cycle control or forcing metabolites away from their optimal levels. It can also be exploited for non-genetic individuality or even for more reliable and deterministic control^{125,37,38}. But even if the pervasive influence of noise has been recognized for decades, it is only recently that quantitative measurements have become feasible. A few pioneering groups have now measured stationary fluctuations over populations of cells, and the results so far are consistent with a few basic principles. The next challenge is to monitor stochastic signals and responses in real time and expose the true cell dynamics buried in population averages. □

Received 8 September; accepted 2 December 2003; doi:10.1038/nature02257.

- Schroedinger, E. *What is Life?* (Cambridge Univ. Press, Cambridge, 1944).
- Novick, A. & Weiner, M. Enzyme induction as an all-or-none phenomenon. *Proc. Natl Acad. Sci. USA* **43**, 553–566 (1957).
- Maloney, P. C. & Rotman, B. Distribution of suboptimally induced β -D-galactosidase in *Escherichia coli*. The enzyme content of individual cells. *J. Mol. Biol.* **73**, 77–91 (1973).
- Löbner-Olesen, A. Distribution of minichromosomes in individual *Escherichia coli* cells: implications for replication control. *EMBO J.* **18**, 1712–1721 (1999).
- Becskei, A. & Serrano, L. Engineering stability in gene networks by autoregulation. *Nature* **405**, 590–593 (2000).
- Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nature Genet.* **31**, 69–73 (2002).
- Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297**, 183–186 (2002).
- Blake, W. J., Kaern, M., Cantor, C. R. & Collins, J. J. Noise in eukaryotic gene expression. *Nature* **422**, 633–637 (2003).
- Xie, X. S. Single-molecule approach to dispersed kinetics and dynamic disorder: Probing conformational fluctuation and enzymatic dynamics. *J. Chem. Phys.* **117**, 11024–11032 (2002).
- Zwanzig, R. *Nonequilibrium Statistical Mechanics* (Oxford Univ. Press, New York, 2001).
- van Kampen, N. G. *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).
- Elf, J., Paulsson, J., Berg, O. G. & Ehrenberg, M. Near-critical phenomena in intracellular metabolite pools. *Biophys. J.* **84**, 154–170 (2003).
- Keizer, K. *Statistical Thermodynamics of Nonequilibrium Processes* (Springer, Berlin, 1987).
- Elf, J. & Ehrenberg, M. Fast evaluations of fluctuations in biochemical networks with the linear noise approximation. *Genome Res.* **13**, 2475–2484 (2003).
- Kacser, H. & Burns, J. A. The control of flux. *Symp. Soc. Exp. Biol.* **27**, 65–104 (1973).
- Westerhoff, H. V. & Van Dam, K. *Thermodynamics and Control of Biological Free Energy Transduction* (Elsevier, Amsterdam, 1987).
- Savageau, M. A. *Biochemical Systems Analysis* (Addison-Wesley, Reading, 1976).
- Savageau, M. A. Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. *Nature* **229**, 542–544 (1971).
- Paulsson, J. & Ehrenberg, M. Noise in a minimal regulatory network: plasmid copy number control. *Q. Rev. Biophys.* **34**, 1–59 (2001).
- Dogterom, M. & Leibler, S. Physical aspects of the growth and regulation of microtubule structures. *Phys. Rev. Lett.* **70**, 1347–1350 (1993).
- Berg, O. G., Paulsson, J. & Ehrenberg, M. Fluctuations and quality of control in biological cells: zero-order ultrasensitivity revisited. *Biophys. J.* **79**, 1228–1236 (2000).
- Rigney, D. R. & Schieve, W. C. Stochastic model of linear, continuous protein-synthesis in bacterial populations. *J. Theor. Biol.* **69**, 761–766 (1977).
- Berg, O. G. A model for statistical fluctuations of protein numbers in a microbial population. *J. Theor. Biol.* **173**, 307–320 (1978).
- McAdams, H. H. & Arkin, A. Stochastic mechanisms in gene expression. *Proc. Natl Acad. Sci. USA* **94**, 814–819 (1997).
- Paulsson, J., Berg, O. G. & Ehrenberg, M. Stochastic focusing: fluctuation-enhanced sensitivity of intracellular regulation. *Proc. Natl Acad. Sci. USA* **97**, 7148–7153 (2000).
- Thattai, M. & van Oudenaarden, A. Intrinsic noise in gene regulatory networks. *Proc. Natl Acad. Sci. USA* **98**, 8614–8619 (2001).
- Peccoud, J. & Ycart, B. Markovian modelling of gene-product synthesis. *Theor. Popul. Biol.* **48**, 222–234 (1995).
- Kepler, T. B. & Elston, T. C. Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. *Biophys. J.* **81**, 3116–3136 (2001).
- Sasai, M. & Wolynes, P. G. Stochastic gene expression as a many-body problem. *Proc. Natl Acad. Sci. USA* **100**, 2374–2379 (2003).
- Singh, U. N. Polyribosomes and unstable messenger RNA: a stochastic model of protein synthesis. *J. Theor. Biol.* **25**, 444–460 (1969).
- Swain, P. S., Elowitz, M. B. & Siggia, E. D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. Natl Acad. Sci. USA* **99**, 12795–12800 (2002).
- Savageau, M. A. Comparison of classical and autogenous systems of regulation in inducible operons. *Nature* **252**, 546–549 (1974).
- Thieffry, D., Huerta, A. M., Perez-Rueda, E. & Collado-Vides, J. From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in *Escherichia coli*. *Bioessays* **20**, 433–440 (1998).
- Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genet.* **31**, 64–68 (2002).
- Paulsson, J. Multilevel selection on plasmid replication. *Genetics* **161**, 1373–1384 (2002).
- Chattoraj, D. K. Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol. Microbiol.* **37**, 467–476 (2000).
- Vilar, J. M., Kueh, H. Y., Barkai, N. & Leibler, S. Mechanisms of noise-resistance in genetic oscillators. *Proc. Natl Acad. Sci. USA* **99**, 5988–5992 (2002).
- Howard, M. & Rutenberg, A. D. Pattern formation inside bacteria: fluctuations due to the low copy number of proteins. *Phys. Rev. Lett.* **90**, 128102 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements This work was supported by a Lewis-Thomas fellowship from Princeton University and by the Cambridge-MIT Institute. I thank A. Becskei, E. Cox, M. Ehrenberg, J. Elf, M. Elowitz, I. Golding, S. Leibler, A. Löbner-Olesen, C. Peterson, S. Sawai, P. Swain, M. Thattai and A. van Oudenaarden for discussions and comments on the manuscript. I also thank W. Blake, M. Kaern, C. Cantor and J. Collins for sharing unpublished data.

Competing interests statement The author declares that he has no competing financial interests.

Correspondence and requests for materials should be addressed to J.P. (J.Paulsson@damtp.cam.ac.uk).

fgf8 mRNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo

Julien Dubrulle & Olivier Pourqu  

Stowers Institute for Medical Research, 1000 East 50th Street, 64110 Kansas City, Missouri, USA

Formation and patterning of the vertebrate embryo occur in a head-to-tail sequence. This progressive mode of body formation from the posterior end of the embryo requires a strict temporal coordination of tissue differentiation—a process involving fibroblast growth factor (FGF) signalling. Here we show that transcription of *fgf8* messenger RNA is restricted to the growing posterior tip of the embryo. *fgf8* mRNA is progressively degraded in the newly formed tissues, resulting in the formation of an mRNA gradient in the posterior part of the embryo. This *fgf8* mRNA gradient is translated into a gradient of FGF8 protein, which correlates with graded phosphorylation of the kinase Akt, a downstream effector of FGF signalling. Such a mechanism provides an efficient means to monitor the timing of FGF signalling, coupling the differentiation of embryonic tissues to the posterior elongation of the embryo. In addition, this mechanism provides a novel model for morphogen gradient formation.

Posterior elongation of the vertebrate embryonic axis is driven initially by the blastopore closure or the regression of the node along the primitive streak and subsequently by the caudal movement of the tail bud. This particular developmental strategy gives rise to an apparent rostral-to-caudal wave of differentiation, whereby tissues located anteriorly are more advanced in their development than are tissues located posteriorly. This progression is well illustrated by the sequential production of somites from the paraxial mesoderm or by the rostro-caudal emigration wave of the neural crest from the neural tube.

Cells in the posterior-most tissues have been proposed to be maintained undifferentiated by a high level of FGF signalling and to activate their differentiation program only when they reach the appropriate threshold of FGF activity as they become located more distantly to the tail bud¹. Much of this activity seems to be mediated by FGF8, which controls activation of the segmentation process in the presomitic mesoderm (PSM)^{2,3} and onset of differentiation in the neural tube^{4–6}. Because of the posterior elongation of the embryo, such a system implies a tight temporal control of FGF signalling in the newly formed posterior tissues. This requires that the highest activity of FGF signalling is constantly maintained in the newly produced cells of the posterior-most end of the embryo. In chick, *fgf8* transcripts show a graded distribution in the PSM and the adjacent neural tube, extending from the tail-bud area to the rostral third of the unsegmented region³.

Graded distribution of FGF8 in the posterior embryo

Here, we first investigated whether the *fgf8* mRNA gradient can control the graded activation of the FGF pathway in the posterior embryo. *In situ* hybridization of mouse embryos with an *fgf8* probe indicated that, as in the chick embryo, *fgf8* is expressed in a graded fashion in the three germ layers (Fig. 1a–d, f–j). This suggests that the posterior *fgf8* mRNA gradient is a conserved feature among vertebrates. Examination of the cellular distribution of *fgf8* mRNA in the PSM gradient in chick and mouse embryos that had been stained for a short period of time clearly showed a salt-and-pepper expression pattern, indicating that the gradient is not smooth (Fig. 1e–j and Supplementary Fig. 1).

We used real-time polymerase chain reaction with reverse transcription (RT–PCR) to characterize further the shape of the *fgf8*

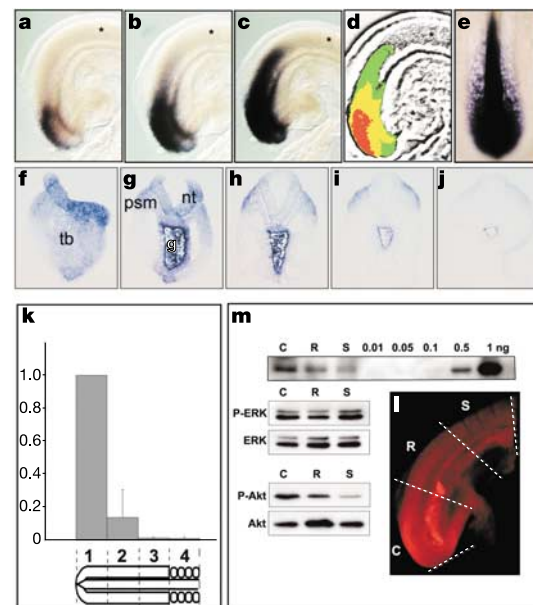


Figure 1 Characterization of the posterior gradient of *fgf8* mRNA and FGF8 protein. **a–c**, Tail region of the same E9.0 mouse embryo hybridized with an *fgf8* probe and photographed after increasing time in the staining reaction: 2 h (**a**), 4 h (**b**), 9 h (**c**). **d**, Overlay of the *fgf8* expression domains shown in **a–c**. **e**, Caudal region of a 2-day-old chick embryo hybridized with an *fgf8* probe after short development of the staining reaction to detect the salt-and-pepper expression pattern in the PSM. **f–j**, Progressively more anterior transverse cryosections of an E9.5 mouse embryo tail hybridized with an *fgf8* probe. **k**, Quantitative analysis of the amount of *fgf8* mRNA by real-time PCR in the four regions of the embryo indicated in the scheme below. **l**, Whole-mount immunohistochemistry on an E9.0 mouse tail probed with an antibody against FGF8. Hatched lines separate the three regions (C, R, S) of the posterior part of the embryo used for the biochemical analysis shown in **m**. **m**, Top, western blot of FGF8 in the regions indicated in **l**. The right part of the blot shows a concentration range (0.01–1 ng) of purified recombinant FGF8b. Bottom, phosphorylation state of ERK and Akt in the regions indicated in **l**. P-ERK and P-Akt denote the phosphorylated isoforms of the proteins. tb, tail bud; psm, presomitic mesoderm; nt, neural tube; g, hindgut. Asterisks in **a–c** indicate the last formed somite.

mRNA gradient in the mouse embryo. Four regions of equivalent size along the antero-posterior (AP) axis were delimited from the trunk of mouse embryos at embryonic day 9 (E9) to E10: three in the PSM, and one including the adjacent newly formed somites (Fig. 1k, bottom). The variation in the *fgf8* content of the four regions was then analysed by real-time RT-PCR and normalized to *Gapdh* (Fig. 1k, top). This analysis showed that the concentration of *fgf8* mRNA experiences a sharp drop-off in the rostral-PSM, indicating that the *fgf8* mRNA gradient in the caudal part of the embryo is not linear.

Depending on the rate of protein translation and on the stability and diffusion properties of the mature secreted protein, the final distribution of FGF8 protein need not necessarily be similar to that of its mRNA. To address this issue, we compared the distribution of *fgf8* mRNA and FGF8 protein in the mouse embryo. We first analysed the distribution of FGF8 protein along the AP axis of the murine embryo by immunohistochemical assessment of whole-mount embryos or parasagittal sections using a monoclonal antibody against the mouse FGF8 protein. This analysis showed that the protein is abundant in the tail bud and that its expression diminishes progressively caudo-rostrally in all three germ layers (Fig. 1l and data not shown).

Expression of FGF8 was assessed quantitatively by western blot analysis. Three regions of equivalent size along the AP axis were delimited from the trunk of E9–E10 mouse embryos (Fig. 1l): a region including the caudal half of the PSM (C), a region including the rostral half of the PSM (R), and a region including the adjacent newly formed somites (S). Protein extracts were prepared from pools of the three different regions described above and analysed. The FGF8 signal decreased in intensity in more rostral tissues (Fig. 1m).

We quantified the total amount of FGF8 protein in these three different regions of the embryo and found that the FGF8 concentration was of the order of 500 ng ml^{-1} (45 nM) in the caudal fragment. The FGF8 concentration progressively diminished rostrally to values in the somitic region that were between two- and fivefold lower than in the caudal-most region. These results indicate that the *fgf8* mRNA gradient is translated into a shallower protein gradient spanning a similar domain of the posterior mouse embryo. Given that the ED_{50} of FGF8b in a mitogenic assay linked to the number of cell divisions experienced by quiescent NR6R-3T3 cells is about $1\text{--}3 \text{ ng ml}^{-1}$ and the dissociation constant (K_d) for the high-affinity FGF receptors is about 10^{-11} M , the observed concentrations in the caudal part of the embryo are within the

physiological range and are consistent with a gradient of FGF8 activity in the PSM⁷.

To examine this possibility, we monitored the phosphorylation state of extracellular-signal-regulated kinase (ERK) and Akt, two kinases involved in mediating the intracellular response downstream of FGF signalling through the Ras-dependent and phosphatidylinositol-3-OH kinase (PI(3)K)-dependent pathways, respectively⁸. No clear graded response of the phosphorylated isoforms of ERK was observed in mouse, consistent with immunocytochemical observations in the mouse primitive streak and PSM⁹. By contrast, the phosphorylated form of the Akt kinase clearly showed a graded profile with a strong signal in the unsegmented region, followed by a sharp decrease in the segmented region (Fig. 1m). These results are therefore consistent with the idea that the *fgf8* mRNA gradient is translated into a ligand gradient, which then controls a graded response along the AP axis in the caudal part of the embryo.

Gradient formation by *fgf8* mRNA decay

We examined how this gradient of mRNA is established and regulated during the course of axis elongation. The existence of a signal emitted by the tail end of the embryo that could activate *fgf8* transcription in a dose-dependent fashion would provide a plausible mechanism of linking axis regression and gradient formation. To test this hypothesis, we compared the expression of *fgf8* in explants of the posterior half of chick embryos with similar

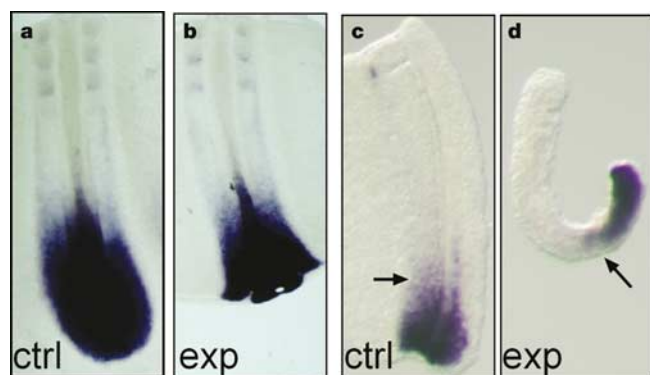


Figure 2 Expression of *fgf8* is not regulated by extrinsic signals. **a, b**, Posterior half explants of 17-somite stage chick embryos hybridized with an *fgf8* probe after a culture period of 210 min. **a**, Intact explant (ctrl); **b**, tail-bud ablation (exp). **c, d**, Intact left posterior half of a 2-day chick embryo (**c**) and contralateral isolated PSM (**d**), both cultured for 105 min and hybridized with an *fgf8* probe. Arrow, the rostral limit of *fgf8* expression in the PSM.

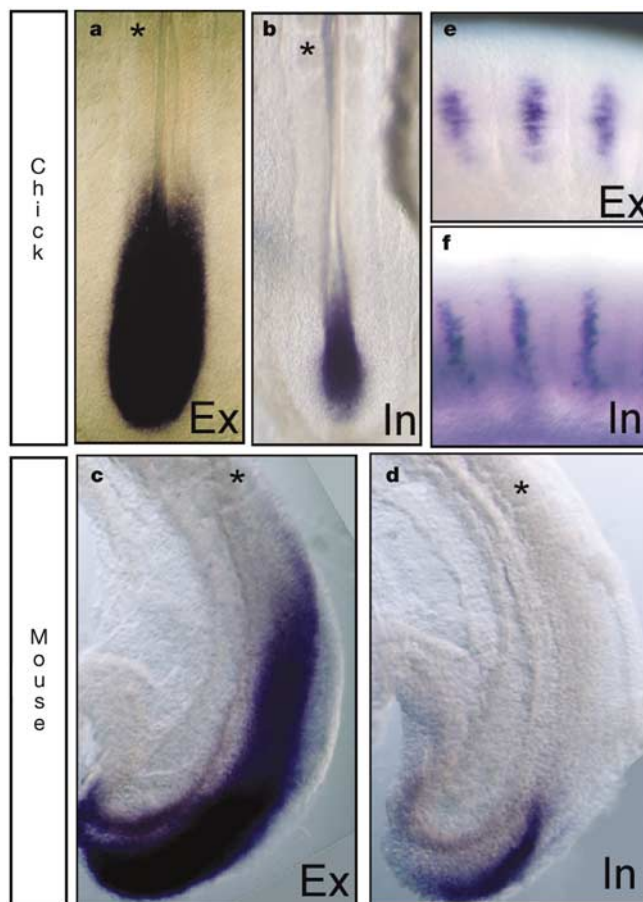


Figure 3 *fgf8* mRNA is only transcribed at the tail-bud level of mouse and chick embryos. Shown is a comparison of exonic (Ex) and intronic (In) *fgf8* expression in chick (**a, b, e, f**) and mouse (**c, d**) embryos. **a, b**, Embryonic tail region in 2-day-old chick embryos. **c, d**, Embryonic tail region in E9.5 mouse embryos. Asterisk, the last formed somite. **e, f**, Labelled myotomes from the trunk region in 3-day-old chick embryos.

explants in which the tail-bud region had been surgically removed (Fig. 2a, b). In the operated explants, the kinetics of the *fgf8* gradient regression was similar to that in the control explants, arguing against the requirement for a tail-bud-derived signal.

We further compared the regression of the *fgf8* mRNA gradient between explant cultures of isolated PSM dissected from one side of 2-day-old chick embryos with their intact contralateral half (Fig. 2c, d). After different culture periods, the intensity and rostral boundary of the *fgf8* expression domain were found to be similar between the dissected tissues and the intact embryonic halves, indicating that establishment of the gradient does not require signalling from the adjacent structures and is therefore an autonomous property of the caudal tissues.

An alternative mechanism based on progressive RNA decay could also account for the autonomous establishment of the caudal *fgf8* mRNA gradient. If *fgf8* mRNA transcription is restricted to the progenitor area in the tail bud and stops when cells become incorporated into the caudal-forming tissues derived from this area, then, taking into account the RNA degradation in these tissues, cells located more rostrally will contain less RNA than will cells that are located in the caudal part of the newly formed tissues (which have been more recently produced by the tail bud). To test this hypothesis, we monitored the production of *fgf8* nascent transcripts in chick and mouse embryos using riboprobes directed against intronic sequences of the *fgf8* gene (Fig. 3).

Whereas *fgf8* expression detected by the exonic probe was observed up to the rostral third of the unsegmented region in chicken embryos at the 20-somite stage (Fig. 3a), cells recognized by the intronic probe were strictly localized to the areas containing the paraxial mesoderm and neural tube progenitors of the tail bud (ref. 10 and Fig. 3b). A similar restriction of active *fgf8* transcription in the tail bud was observed in the mouse embryo (Fig. 3c, d). This difference in expression is not due to lower sensitivity of the intronic probe, because intronic expression was observed in virtually all known expression sites of exonic *fgf8*, such as the midbrain–hindbrain boundary, the myotome and the apical ectodermal ridge (AER) (Fig. 3f and data not shown). Nuclear expression of the intronic transcripts was particularly obvious in the chick myotome, where nuclei are aligned as a single stripe in the middle of the structure (compare Fig. 3f and e). These observations therefore suggest that the *fgf8* mRNA gradient seen in the caudal embryo does not result from active transcription but from the decay of mRNA transcribed at the level of the tail-bud progenitors.

To test this hypothesis further, we cultured the caudal portion of chick embryos at the 18–23-somite stage *in vitro* in medium containing the transcription inhibitor actinomycin D¹¹. We then analysed the expression of mature *fgf8* transcripts by *in situ*

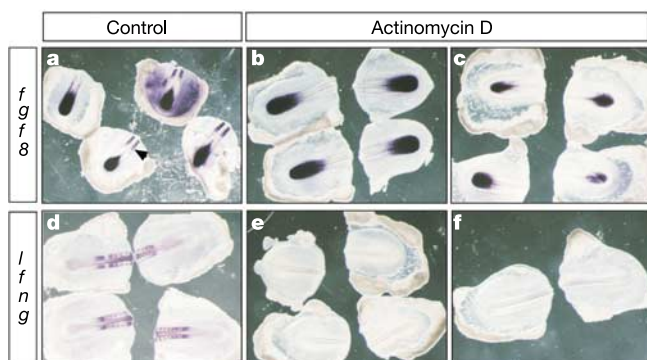


Figure 4 High stability of processed *fgf8* mRNAs. **a–f**, Posterior halves of chick embryos hybridized with exonic *fgf8* (**a–c**) or *lfrng* (**d–f**) after *in vitro* culture in the absence (control) or presence of actinomycin D. **a, c, d, f**, After 5 h of culture; **b, e**, after 3 h of culture.

hybridization after different times in culture. As a control, we monitored expression of the mature transcripts of *lunatic fringe* (*lfrng*), a gene that has a rapid turnover of mRNA in the PSM under the same culture conditions¹². Expression of *fgf8* transcripts in caudal tissues was still present and distributed in a graded fashion after incubation in actinomycin-D-containing medium for 3–5 h (Fig. 4a–c), whereas expression of *lfrng* was totally abolished from embryos treated in the same conditions (Fig. 4d–f). Efficiency of the treatment was confirmed further by the absence of the *fgf8* expression that was initiated at the time of somite formation and later found in the posterior part of the somites in the treated embryos (Fig. 4a, arrow).

These results suggest that the *fgf8* transcripts detected in the caudal cells of the embryo did not derive from active transcription but instead were originally produced in the tail bud and inherited by their descendants in the PSM and the neural tube. The salt-and-pepper nature of *fgf8* mRNA within the gradient (Fig. 1e) is consistent with local cellular rearrangements that are known to occur in the posterior part of the embryo¹³.

Discussion

Few cell divisions are observed in the PSM, suggesting that *fgf8* RNA decay along the AP axis is unlikely to result from massive clonal dilution¹⁴. Considering a constant rate of translation, therefore, the slope of the FGF8 protein gradient will be essentially controlled by the stability of the transcripts. Combined with the continuous production of committed cells caudally, these molecular mecha-

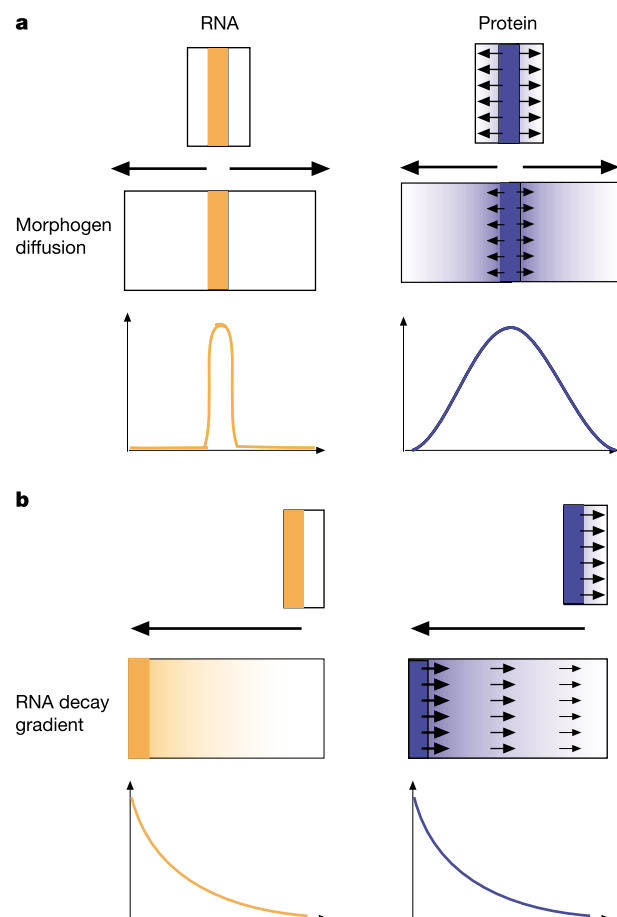


Figure 5 Two different mechanisms of establishing a protein gradient. **a**, Gradient formation by protein diffusion from a localized source. **b**, Gradient formation through polarized growth and mRNA decay.

nisms are indeed theoretically capable of generating a graded, dynamic distribution of a secreted molecule during elongation of the axis of the vertebrate embryo.

The establishment of gradients of secreted proteins during development have been studied mainly for morphogens such transforming growth factor- β or Wnt family members¹⁵. The formation of these gradients is described for systems undergoing even growth, and the gradients rely on the existence of a localized source producing the morphogen in the same place where the RNA transcription of the morphogen occurs (Fig. 5a).

The gradient is then established from the source across the developing field by specific diffusion properties of the protein or by active transport mechanism.

Our results provide evidence for a previously unrecognized mode of gradient formation in the embryo. In this model, the protein gradient is moulded onto the mRNA gradient, which is established owing to the peculiar polarized growth mode of the vertebrate axis (Fig. 5b). Such a patterning strategy of the embryonic axis might be used widely by other animals, such as invertebrates that show a sequential mode of development¹⁶. It might also be involved in the development of other embryonic structures arising by polarized growth, such as those derived from epithelial cords, as is seen in the development of many internal organs¹⁷. □

Methods

Plasmids and *in situ* hybridization

Whole-mount *in situ* hybridizations were done as described¹⁸. To generate intronic chick and mouse *fgf8* riboprobes, we first amplified the fourth intron of the *fgf8* gene by PCR on genomic DNA using the primers F (5'-GGAAATTCATCTGCATGAACAAGAAGG-3') and R (5'-GGAATTCACGCAGTCCCTTGCCCTTGGCCG-3'), and then cloned it into pGEMTeasy (Promega). The mouse fragment was further subcloned into pBluescript (Stratagene) after a *Pst*I-*Sac*I digestion to remove all flanking exonic sequences. The chick and mouse exonic *fgf8* and the *lfrg* riboprobes have been described^{12,19,20}.

Real-time RT-PCR

Real-time RT-PCR analysis was done with the iCycler iQ detection system (Bio-Rad). In brief, E9.0 embryonic tails were cut in four equal parts (see Fig. 1k), pooled in groups of two, and subjected to total RNA extraction by the RNeasy minikit (Qiagen). Each RNA sample was divided into two and subjected to RT-PCR for *fgf8* and *Gadph* within the same run using Quantitect SYBR Green buffer (Qiagen). We used the following primers: for *fgf8*, F (5'-CAGCTCTACAGCCGACACAGC-3') and R (5'-TGCTCTTGCAATTAGC TCC-3'); for *Gadph*, F (5'-AACGACCCCTTCATTGAC-3') and R (5'-TCCACGACATAC TCAGCAC-3').

Before real-time RT-PCR, the quality and purity of amplified fragments, as well as protocol optimization, were checked after traditional RT-PCR by agarose gel electrophoresis. Results are presented as the fold expression of *fgf8* relative to the caudal-most part of the tail, which was calculated using 2^{-ddC_t} , where $ddC_t = fgf8 C_t - Gadph C_t$ and $ddC_t = dC_t$ of the caudal part - dC_t of a given region²¹. The C_t value is the calculated threshold cycle number at which the fluorescent signal is significantly above the background. The graph in Fig. 1k plots the mean $\pm$ s.d. of the fold expression of each of the four regions from a total of six embryos.

FGF8 protein detection

Immunohistochemistry was done by standard procedures using a monoclonal antibody against FGF8b (R&D Systems), a biotinylated antibody against mouse IgG1 (Southern Biotech) and Cy3-conjugated streptavidin (Sigma).

Western blot analysis was done by standard procedures. The caudal region of mouse embryos at the 20–30-somite stage was divided into three equal parts (see Fig. 1l) corresponding to the caudal half, the rostral half of the unsegmented region and a region including the last formed somites. We lysed samples in RIPA buffer containing 1 mM NaF, 1 mM Na_2VO_4 , 1 mM sodium pyrophosphate and Complete proteases inhibitor cocktail (Roche). Proteins were extracted from pools of the different regions and separated by SDS-PAGE on 12% acrylamide gels. The concentration of FGF8 in tissue was evaluated by comparing signal intensity in the embryo samples (normalized to β -actin) to known dilutions of recombinant mouse FGF8b (R&D Systems).

The FGF8 signal intensity was quantified either from the autoradiogram after detection by electrochemiluminescence (Amersham) using Image software (NIH) and normalization to the expression of β -actin (A-2066, Sigma) ($n = 8$), or by western blotting using Alexa-coupled secondary antibodies and a Li-cor scanning system ($n = 3$).

Both methods gave similar estimations. Antibodies against ERK (9101), phosphorylated ERK (9102), Akt (9272) and phosphorylated Akt (9271) were from Cell Signalling Technology, and western blots were done in accordance with the manufacturer's instructions. Western blots were reproduced at least three times for each antibody.

In vitro chick embryo culture and actinomycin D treatment

Chick embryo explants at the 18–23-somite stage were cultured *in vitro* as described³. In a first series of experiments, intact posterior half explants and posterior half explants where the tail-bud region was surgically removed were cultured in parallel for up to 230 min. In a second series, isolated PSM and contralateral intact half embryos were cocultured for 60–270 min. In a third series, intact posterior half embryos were cultured in the presence or absence of 5 or 10 μ g ml⁻¹ actinomycin D (Sigma) for up to 5 h. We then fixed the explants for *in situ* hybridization.

Received 7 July; accepted 5 November 2003; doi:10.1038/nature02216.

- Vasiliou, D. & Stern, C. D. Patterning the embryonic axis: FGF signaling and how vertebrate embryos measure time. *Cell* **106**, 133–136 (2001).
- Sawada, A. *et al.* Fgf/MAPK signalling is a crucial positional cue in somite boundary formation. *Development* **128**, 4873–4880 (2001).
- Dubrulle, J., McGrew, M. J. & Pourqu  , O. FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* **106**, 219–232 (2001).
- Mathis, L., Kulesa, P. M. & Fraser, S. E. FGF receptor signalling is required to maintain neural progenitors during Hensen's node progression. *Nature Cell Biol.* **3**, 559–566 (2001).
- Diez del Corral, R., Breitkreuz, D. N. & Storey, K. G. Onset of neuronal differentiation is regulated by paraxial mesoderm and requires attenuation of FGF signalling. *Development* **129**, 1681–1691 (2002).
- Bertrand, N., Medevielle, F. & Pituello, F. FGF signalling controls the timing of Pax6 activation in the neural tube. *Development* **127**, 4837–4843 (2000).
- Chelliah, A., Yuan, W., Chelliah, M. & Ornitz, D. M. Mapping ligand binding domains in chimeric fibroblast growth factor receptor molecules. Multiple regions determine ligand binding specificity. *J. Biol. Chem.* **274**, 34785–34794 (1999).
- Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **103**, 211–225 (2000).
- Corson, L. B., Yamanaka, Y., Lai, K. M. & Rossant, J. Spatial and temporal patterns of ERK signaling during mouse embryogenesis. *Development* **130**, 4527–4537 (2003).
- Catala, M., Teillet, M. A. & Le Douarin, N. M. Organization and development of the tail bud analyzed with the quail–chick chimera system. *Mech. Dev.* **51**, 51–65 (1995).
- Hyman, R. W. & Davidson, N. Kinetics of the *in vitro* inhibition of transcription by actinomycin. *J. Mol. Biol.* **50**, 421–438 (1970).
- McGrew, M. J., Dale, J. K., Fraboulet, S. & Pourqu  , O. The *lunatic fringe* gene is a target of the molecular clock linked to somite segmentation in avian embryos. *Curr. Biol.* **8**, 979–982 (1998).
- Kulesa, P. M. & Fraser, S. E. Cell dynamics during somite boundary formation revealed by time-lapse analysis. *Science* **298**, 991–995 (2002).
- Stern, C. D., Fraser, S. E., Keynes, R. J. & Primmitt, D. R. A cell lineage analysis of segmentation in the chick embryo. *Development* **104**, 231–244 (1988).
- Gurdon, J. B. & Bourillot, P. Y. Morphogen gradient interpretation. *Nature* **413**, 797–803 (2001).
- Davis, G. K. & Patel, N. H. The origin and evolution of segmentation. *Trends Cell Biol.* **9**, M68–M72 (1999).
- Hogan, B. L. & Kolodziej, P. A. Organogenesis: molecular mechanisms of tubulogenesis. *Nature Rev. Genet.* **3**, 513–523 (2002).
- Henrique, D. *et al.* Expression of a Delta homologue in prospective neurons in the chick. *Nature* **375**, 787–790 (1995).
- Crossley, P. H. & Martin, G. R. The mouse *Fgf8* gene encodes a family of polypeptides and is expressed in regions that direct outgrowth and patterning in the developing embryo. *Development* **121**, 439–451 (1995).
- Crossley, P. H., Minowada, G., MacArthur, C. A. & Martin, G. R. Roles for FGF8 in the induction, initiation, and maintenance of chick limb development. *Cell* **84**, 127–136 (1996).
- Bottone, F. G. Jr, Martinez, J. M., Collins, J. B., Afshari, C. A. & Eling, T. E. Gene modulation by the cyclooxygenase inhibitor, sulindac sulfide, in human colorectal carcinoma cells: possible link to apoptosis. *J. Biol. Chem.* **278**, 25790–25801 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank D. Duboule for discussions; G. Martin for reagents; B. Brede for help with real-time PCR; S. Fraser, D. Ish-Horowitz, P. Kulesa, T. Lecuit and members of the Pourqu   laboratory for comments on the manuscript. J.D. was a recipient of a fellowship from the Fondation pour la Recherche M  dicale (FRM). This work was initiated in the Laboratoire de G  n  tique et Physiologie du D  veloppement in Marseille, supported by funding from the Centre National de la Recherche Scientifique (CNRS), Human Frontier Science Program Organization (HFSPO), Association Fran  aise contre les Myopathies and the Universit   de la M  diterran  e-AP de Marseille. Current work is supported by the Stowers Institute and a grant from the NIH.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to O.P. (olp@stowers-institute.org).

A high-mobility electron gas at the $\text{LaAlO}_3/\text{SrTiO}_3$ heterointerface

A. Ohtomo^{1,2,3} & H. Y. Hwang^{1,3,4}

¹Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA

²Institute for Materials Research, Tohoku University, Sendai, 980-8577, Japan

³Japan Science and Technology Agency, Kawaguchi, 332-0012, Japan

⁴Department of Advanced Materials Science, University of Tokyo, Kashiwa, Chiba, 277-8651, Japan

Polarity discontinuities at the interfaces between different crystalline materials (heterointerfaces) can lead to nontrivial local atomic and electronic structure, owing to the presence of dangling bonds and incomplete atomic coordinations^{1–3}. These discontinuities often arise in naturally layered oxide structures, such as the superconducting copper oxides and ferroelectric titanates, as well as in artificial thin film oxide heterostructures such as manganite tunnel junctions^{4–6}. If polarity discontinuities can be atomically controlled, unusual charge states that are inaccessible in bulk materials could be realized. Here we have examined a model interface between two insulating perovskite oxides— LaAlO_3 and SrTiO_3 —in which we control the termination layer at the interface on an atomic scale. In the simple ionic

limit, this interface presents an extra half electron or hole per two-dimensional unit cell, depending on the structure of the interface. The hole-doped interface is found to be insulating, whereas the electron-doped interface is conducting, with extremely high carrier mobility exceeding $10,000 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1}$. At low temperature, dramatic magnetoresistance oscillations periodic with the inverse magnetic field are observed, indicating quantum transport. These results present a broad opportunity to tailor low-dimensional charge states by atomically engineered oxide heteroepitaxy.

An early discussion of polarity or valence discontinuities arose in the consideration of the growth of GaAs on (001)-oriented Ge^{1,2}. Both semiconductors have the same crystal structure and nearly exact lattice match, thus representing promising materials to combine direct and indirect bandgap semiconductor functions. Just at the interface, however, there are incomplete bonds at the termination of the group IV Ge layer and the commencement of III–V alternations of GaAs. There have been recent attempts to design interfaces on the atomic scale to compensate for these dangling bonds⁷. Layered oxide crystal structures can be viewed as an intimate sequence of valence discontinuities, often involving charge-transfer over a few atomic positions. The myriad of stacking sequences such as the perovskite-derived Ruddlesden–Popper phases, constructed as $A_{n+1}B_nO_{3n+1}$, for $0 \leq n \leq \infty$, involve accommodating this charge transfer while maintaining global charge neutrality^{8,9}. Recently, lamellar contacts between members

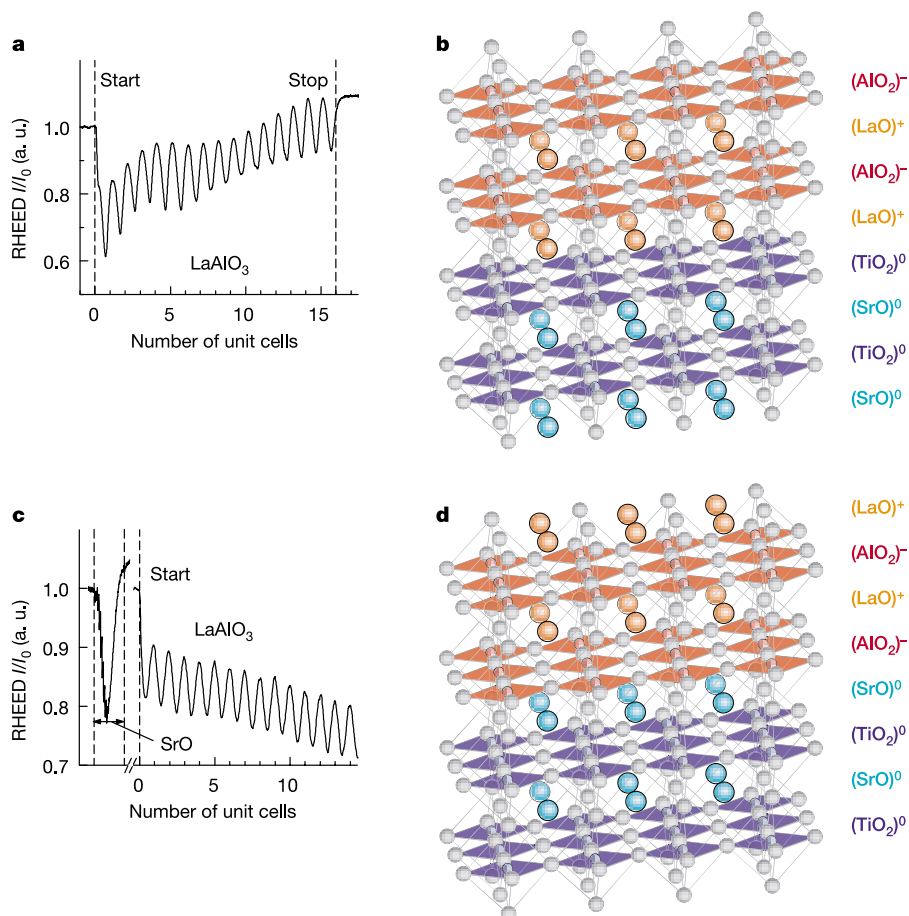


Figure 1 Growth and schematic models of the two possible interfaces between LaAlO_3 and SrTiO_3 in the (001) orientation. **a**, RHEED intensity oscillations of the specular reflected beam for the growth of LaAlO_3 directly on the TiO_2 terminated SrTiO_3 (001) surface. **b**, Schematic of the resulting $(\text{LaO})^+ / (\text{TiO}_2)^0$ interface, showing the composition

of each layer and the ionic charge state of each layer. **c**, RHEED oscillations for the growth of LaAlO_3 , after a monolayer of SrO was deposited on the TiO_2 surface. **d**, Schematic of the resulting $(\text{AlO}_2)^- / (\text{SrO})^0$ interface.

of the ilmenite–haematite mineral series have been shown to induce interfacial magnetism via charge transfer¹⁰.

Valence discontinuities naturally arise as a general concern in complex oxide heteroepitaxy as well. The electronic structure of an oxide heterointerface is important for its stability and function in many devices. In manganite tunnel junctions, for example, the magnetoresistance of the device is exponentially sensitive to the charge and spin state at the manganite–tunnel barrier interface⁶. The heterointerface between LaAlO_3 and SrTiO_3 is a simple, experimentally accessible realization of a valence discontinuity in perovskite oxides. Both are wide-bandgap insulators ($\text{LaAlO}_3 \approx 5.6$ eV, $\text{SrTiO}_3 \approx 3.2$ eV), and they are reasonably well lattice matched to one another (3.789 Å and 3.905 Å, respectively). The formal valence states can be assigned as La^{3+} , Al^{3+} , O^{2-} , Sr^{2+} , and Ti^{4+} ; to first order, only Ti has accessible mixed-valence character, allowing for reduction towards Ti^{3+} . In the (001) direction, the perovskite structure ABO_3 can be considered as alternating stacks of AO and BO_2 . SrTiO_3 is a sequence of charge-neutral sheets, whereas LaAlO_3 alternates between $\pm e$ -charged sheets in the ionic limit, where e is the electron charge. The heterointerface of LaAlO_3 and SrTiO_3 , therefore, presents an extra $\pm e/2$ per two-dimensional unit cell, again in the simple ionic limit (Fig. 1b and d).

The $\text{LaAlO}_3/\text{SrTiO}_3$ interfaces were grown in an ultrahigh vacuum chamber (Pascal) by pulsed laser deposition using a single-crystal LaAlO_3 target on (001) SrTiO_3 single-crystal substrates. SrTiO_3 substrates are well suited for this study, as techniques have been developed to etch TiO_2 -terminated surfaces characterized by atomically flat terraces separated by unit cell steps¹¹. The $(\text{LaO})^+ / (\text{TiO}_2)^0$ interface (Fig. 1b) was grown by direct deposition on this surface using a KrF excimer laser at a repetition rate of 2 Hz and a laser fluence at the target surface of $\sim 1 \text{ J cm}^{-2}$. The growth temperature was 800 °C as measured by a pyrometer (emissivity 0.8), with an oxygen partial pressure p_{O_2} ranging from 10^{-4} – 10^{-6}

torr. Typical unit cell reflection high-energy electron diffraction (RHEED) intensity oscillations are shown in Fig. 1a. The $(\text{AlO}_2)^- / (\text{SrO})^0$ interface (Fig. 1d) was grown by first depositing a monolayer of SrO from a single-crystal target, then switching to LaAlO_3 *in situ*¹². Typical RHEED oscillations for this case are shown in Fig. 1c. After growth of LaAlO_3 , atomic force microscopy reveals that the original step and terrace structure of the substrate surface is preserved. This indicates that the intended interface is generally formed across the entire substrate.

We examined the conductivity of the interface, which required laser-annealed ohmic contacts through contact shadow masks to reach the buried interface (unpublished work). This contact method produced an array of microcracks penetrating the LaAlO_3 layer, as well as a tub of highly n-type SrTiO_3 formed by oxygen vacancies induced by the radiation. These contacts were formed around transport samples in typical six-probe Hall bar geometries. In all cases studied, the $(\text{AlO}_2)^- / (\text{SrO})^0$ interface, expected to have a p-type interface, was insulating, although the possibility of Schottky barrier formation exists. The $(\text{LaO})^+ / (\text{TiO}_2)^0$ interface, however, indeed displayed n-type conductivity for a wide range of conditions. The temperature-dependent sheet resistance $R_{\text{xx}}(T)$ is given in Fig. 2a and b for 60-Å-thick and 260-Å-thick LaAlO_3 on SrTiO_3 for the $(\text{LaO})^+ / (\text{TiO}_2)^0$ interface at different p_{O_2} values during growth. The thicker LaAlO_3 film showed little p_{O_2} dependence, whereas the thinner film lost significant conductivity for higher p_{O_2} values. For the case of the films grown at 10^{-6} torr, a number of films are shown as a sample of the variability and reproducibility. The temperature dependence of the Hall coefficient R_{H} is given in Fig. 2c and d for the corresponding samples in Fig. 2a and b. There is little or no evidence for carrier freeze-out in most samples, consistent with the low activation barrier for dopants in SrTiO_3 (ref. 13). The resultant Hall mobility $\mu_{\text{H}}(T)$ is given in Fig. 2e and f, demonstrating the extremely high carrier mobility that can be obtained at the interface.

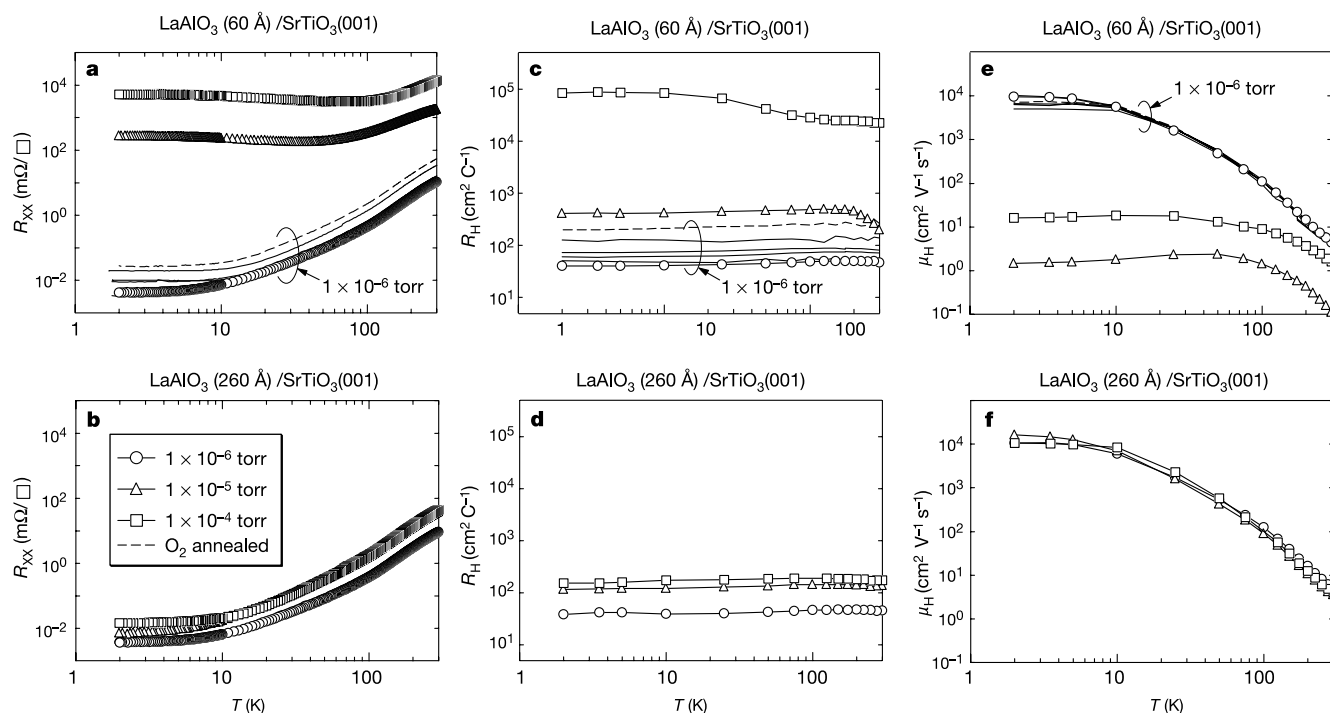


Figure 2 Transport properties of the $(\text{LaO})^+ / (\text{TiO}_2)^0$ interface for different oxygen partial pressures p_{O_2} during growth at 10^{-4} , 10^{-5} , and 10^{-6} torr, as well as for 10^{-6} torr growth followed by annealing in 1 atm of O_2 at 400 °C for 2 h. **a, c, e**, The temperature

dependence of $R_{\text{xx}}(T)$, $R_{\text{H}}(T)$ and $\mu_{\text{H}}(T)$ for the interface between 60-Å-thick LaAlO_3 and SrTiO_3 , respectively. **b, d, f**, The temperature dependence of $R_{\text{xx}}(T)$, $R_{\text{H}}(T)$ and $\mu_{\text{H}}(T)$ for the interface between 260-Å-thick LaAlO_3 and SrTiO_3 , respectively.

One possible origin of the carriers is trapped oxygen vacancies at the surface of the substrate. To rule out this possibility, we confirmed that exposure of the substrate to the growth conditions (800°C , 10^{-6} torr p_{O_2} , held for one hour, then quenched to room temperature) preserves an insulating surface. A related possibility is the formation of oxygen vacancies near the surface of SrTiO_3 by 'gettering' from incompletely oxidized LaAlO_3 in the growing film. This can be ruled out because the conductivity is removed when we insert a single monolayer of SrO before the LaAlO_3 growth. If oxygen gettering is the dominant origin of the carriers, then n-type conductivity should be obtained for both interfaces, which was not observed. Finally, we note that when a fraction of a monolayer of SrO is inserted before LaAlO_3 growth, the carrier density decreases proportionally with increasing SrO coverage from 0 to 1 monolayer. Therefore we conclude that the dominant mechanism introducing

carriers is the built-in polarity discontinuity of the interface.

The high carrier mobilities imply that quantum transport might be observed in moderate magnetic fields at low temperatures. Indeed, distinct magnetoresistance oscillations are observed at 2 K for the 60 Å films grown at 10^{-6} torr, with representative traces given in Fig. 3a for samples shown in Fig. 2. As can be seen, the spacing and magnitude of the response varies widely, sample-to-sample, and sometimes even within the same sample hysteretically in magnetic field H . We have found that these oscillations diminish in magnitude over the course of weeks, until they completely vanish, despite an unchanging $R_{\text{xx}}(T)$ and $\mu_{\text{H}}(T)$. In some cases, however, they are relatively stable, as shown in Fig. 3b for two temperatures. To check the dimensionality of the charge state, we performed rotation experiments in magnetic field, given in Fig. 3c. The magnetoresistance features clearly remain fixed during rotation, indicating three-dimensional behaviour.

These intriguing results introduce a number of unresolved questions. One of the simplest is given in Fig. 4, which shows the variation of $1/R_{\text{H}}e$ for the samples studied here. Interpreting this as the sheet carrier density n would imply unphysical densities (10^{17} cm^{-2} requires a unit cell carrier density of $1.7 \times 10^{22} \text{ cm}^{-3}$ over 600 Å of thickness). Although we cannot rule out a possible role for oxygen vacancies at the interface, this extreme level is unobtainable in our growth conditions, and thermodynamically unfeasible¹⁴. Annealing in 1 atm of O_2 at 400°C for 2 h gives little change to the transport properties. One possible interpretation is that $1/R_{\text{H}}e$ diverges from the simple carrier density in this system, as seen at high doping levels in $(\text{La,Sr})\text{TiO}_3$ (ref. 15). Extreme-high-temperature annealing does diminish the conductivity more dramatically, but we caution that this may be driven by La/Sr interdiffusion. The diffusion at the interface of just the neighbouring sites is sufficient to cancel the interface charge, and the resolution of this Coulomb energy cost may enhance diffusion significantly. At 10^{-4} torr for the 60-Å films, the value associated with $e/2$ per two-dimensional unit cell is approached (Fig. 4).

The absence of the magnetoresistance oscillations in the thicker films is also unaccounted for, despite comparable mobility. Explanations may include kinetic effects involving cation redistribution and oxygen stoichiometry at the interface, or band-bending effects from states on the LaAlO_3 surface¹⁶. In addition, the lineshape of the magnetoresistance oscillation is unusual. Generally, we find one

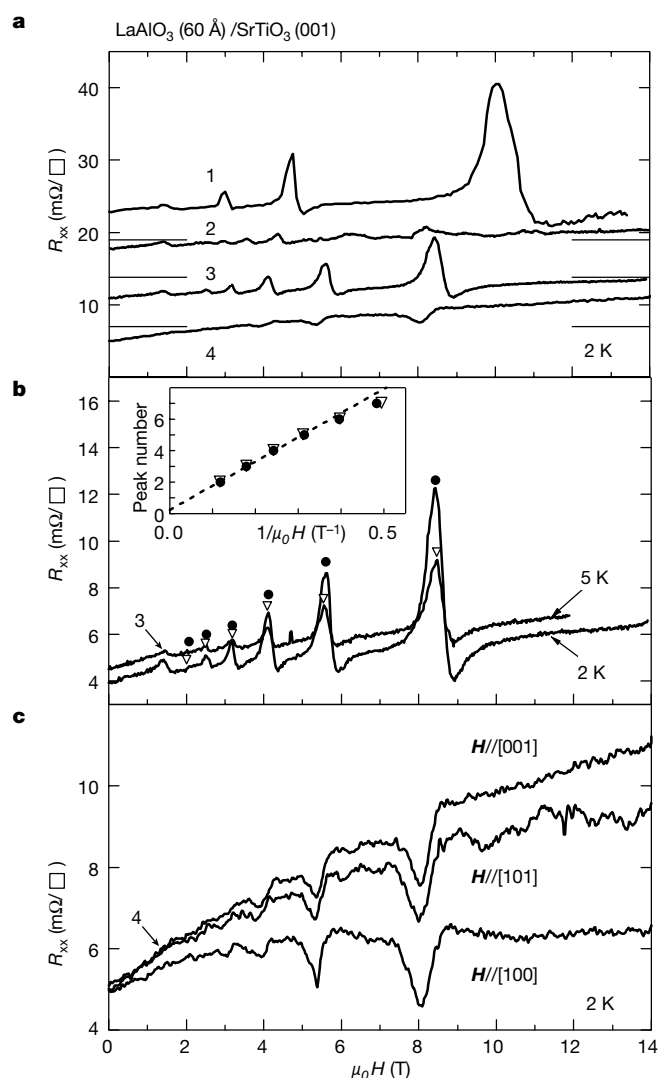


Figure 3 Low-temperature magnetoresistance of the $(\text{LaO})^+ / (\text{TiO}_2)^0$ interface between 60-Å-thick LaAlO_3 and SrTiO_3 grown at 10^{-6} torr p_{O_2} . **a**, Representative transverse magnetoresistance traces taken at 2 K. Curves 1–3 have been offset from the origin for clarity, with the positions of zero denoted by the corresponding horizontal lines. **b**, Magnetoresistance for sample number 3 in **a** at 2 K and 5 K. The peak positions are denoted by dots. Inset, the magnetoresistance peaks plotted versus the inverse magnetic field. **c**, Transverse magnetoresistance for sample number 4 in **a** for different field orientations: perpendicular to the interface ($H//[001]$), at 45° ($H//[101]$), and in the plane ($H//[100]$).

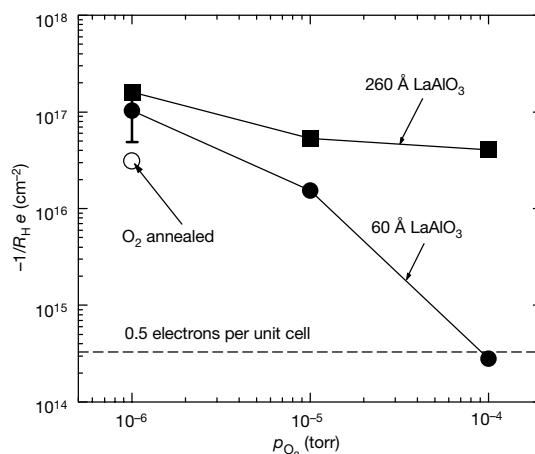


Figure 4 The dependence of $-1/R_{\text{H}}e$ at 2 K on the p_{O_2} during growth for the samples in this study. This quantity gives the carrier density $n = -1/R_{\text{H}}e$ for simple semiconductors. The sheet carrier density corresponding to 0.5 electrons per unit cell is denoted by the dashed line.

principal structure that is periodic in $1/H$. In comparable circumstances for bulk SrTiO_3 , Shubnikov–de Haas oscillations associated with Fermi surface orbits give a sinusoidal variation¹⁷. Rather, we find a structure reminiscent of a Fano lineshape¹⁸.

Despite these open questions, a number of issues can be deduced. Because there is little variation in the $\text{O}2\text{p}$ -derived valence band across the perovskites, most of the band discontinuity must arise in the conduction band, placing the carrier in the SrTiO_3 conduction band. The periodicity in $1/H$ implies that this phenomenon can be associated with orbits in momentum space, rather than the H periodicity associated with real-space orbits¹⁹. These results are consistent with multiple sub-band occupancy arising from the high sheet carrier density, coupled with a very weak confining potential on the SrTiO_3 side due to the large (greater than 10,000), low-temperature static dielectric constant²⁰. The varying behaviour in Fig. 3a is consistent with inhomogeneous current flow, arising from imperfectly uniform termination of the intended $(\text{LaO})^+/(\text{TiO}_2)^0$ type²¹. Friction force microscopy indicates that occasional missed-termination sites can be found, particularly near the step edges. Transport geometries parallel and perpendicular to the step edges show no significant difference, indicating that scattering across step edges does not dominate transport. However, given the typical terrace width of 200 nm, many of these effects are necessarily averaged out in macroscopic transport samples. In this regard, although the magnetoresistance is measured to be macroscopically symmetric in H , mixing of the conductivity tensor may occur on a microscopic scale, especially considering the large Hall angles involved²². Nevertheless, it is a striking result that the periodicity in $1/H$ is unchanged by rotation in the field. This feature is associated with a constant cross-section of the Fermi surface in the standard interpretation, implying an electronic structure very different from that established for bulk SrTiO_3 . We suggest that an unusual, non-bulk-like charge state is formed at this polar heterointerface. □

Received 10 October; accepted 19 December 2003; doi:10.1038/nature02308.

- Baraff, G. A., Appelbaum, J. A. & Hamann, D. R. Self-consistent calculation of the electronic structure at an abrupt GaAs–Ge interface. *Phys. Rev. Lett.* **38**, 237–240 (1977).
- Harrison, W. A., Kraut, E. A., Waldrop, J. R. & Grant, R. W. Polar heterojunction interfaces. *Phys. Rev. B* **18**, 4402–4410 (1978).
- Kroemer, H. Polar-on-nonpolar epitaxy. *J. Cryst. Growth* **81**, 193–204 (1987).
- Bednorz, J. G. & Mueller, K. A. Possible high- T_c superconductivity in the Ba–La–Cu–O system. *Z. Phys. B* **64**, 189–193 (1986).
- Park, B. H. *et al.* Lanthanum-substituted bismuth titanate for use in non-volatile memories. *Nature* **401**, 682–684 (1999).
- Tokura, Y. R. (ed.) *Colossal Magnetoresistive Oxides* Ch. 9, 10 (Gordon and Breach, New York, 2000).
- Wang, T., Moll, N., Cho, K. & Joannopoulos, J. D. Deliberately designed materials for optoelectronic applications. *Phys. Rev. Lett.* **82**, 3304–3307 (1999).
- Ruddlesden, S. N. & Popper, P. New compounds of the K_2NiF_4 type. *Acta Crystallogr.* **10**, 538–540 (1957).
- Ruddlesden, S. N. & Popper, P. The compound $\text{Sr}_3\text{Ti}_2\text{O}_7$ and its structure. *Acta Crystallogr.* **11**, 54–55 (1958).
- Robinson, P., Harrison, R. J., McEnroe, S. A. & Hargraves, R. B. Lamellar magnetism in the haematite–ilmenite series as an explanation for strong remanent magnetization. *Nature* **418**, 517–520 (2002).
- Kawasaki, M. *et al.* Atomic control of the SrTiO_3 crystal-surface. *Science* **266**, 1540–1542 (1994).
- Kim, D. W. *et al.* Roles of the first atomic layers in growth of SrTiO_3 films on LaAlO_3 substrates. *Appl. Phys. Lett.* **74**, 2176–2178 (1999).
- Tufte, O. N. & Chapman, P. W. Electron mobility in semiconducting strontium titanate. *Phys. Rev.* **155**, 796–802 (1967).
- Lide, D. R. (ed.) *CRC Handbook of Chemistry and Physics*, 77th edn (CRC Press, Boca Raton, 1996).
- Tokura, Y. *et al.* Filling dependence of electronic properties on the verge of metal–Mott-insulator transitions in $\text{Sr}_{1-x}\text{La}_x\text{TiO}_3$. *Phys. Rev. Lett.* **70**, 2126–2129 (1993).
- Francis, R. J., Moss, S. C. & Jacobson, A. J. X-ray truncation rod analysis of the reversible temperature-dependent [001] surface structure of LaAlO_3 . *Phys. Rev. B* **64**, 235425 (2001).
- Frederikse, H. P. R., Hosler, W. R. & Thurber, W. R. Shubnikov–de Haas effect in SrTiO_3 . *Phys. Rev.* **158**, 775–778 (1967).
- Fano, U. Effects of configuration interaction on intensities and phase shifts. *Phys. Rev.* **124**, 1866–1878 (1961).
- Pippard, A. B. *Magnetoresistance in Metals* Ch. 6 (Cambridge Univ. Press, Cambridge, 1989).
- Sakudo, T. & Unoki, H. Dielectric properties of SrTiO_3 at low temperatures. *Phys. Rev. Lett.* **26**, 851–853 (1971).
- Nishimura, T. *et al.* Structure change of TiO_2 -terminated $\text{SrTiO}_3(001)$ surfaces by annealing in O_2 atmosphere and ultrahigh vacuum. *Surf. Sci.* **421**, 273–278 (1999).

- Herring, C. Effect of random inhomogeneities on electrical and galvanomagnetic measurements. *J. Appl. Phys.* **31**, 1939–1953 (1960).

Acknowledgements We thank D. Schlom, D. R. Hamann and T. Ohnishi for discussions. We acknowledge partial support from NEDO's International Joint Research Program. A.O. acknowledges partial support from the Asahi Glass Foundation and the Inamori Foundation.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.Y.H. (hyhwang@k.u-tokyo.ac.jp).

Atomic-scale imaging of carbon nanofibre growth

Stig Helveg¹, Carlos López-Cartes^{1*}, Jens Sehested¹, Poul L. Hansen¹, Bjerne S. Clausen¹, Jens R. Rostrup-Nielsen¹, Frank Abild-Pedersen² & Jens K. Nørskov²

¹Haldor Topsøe A/S, Nymøllevej 55, and ²Center for Atomic-scale Materials Physics, Department of Physics, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

* Present address: Instituto de Ciencia de Materiales de Sevilla (CSIC-UNSE), Avda. Américo Vespucio s/n, Isla de la Cartuja, 41092 Sevilla, Spain

The synthesis of carbon nanotubes with predefined structure and functionality plays a central role in the field of nanotechnology^{1,2}, whereas the inhibition of carbon growth is needed to prevent a breakdown of industrial catalysts for hydrogen and synthesis gas production³. The growth of carbon nanotubes and nanofibres has therefore been widely studied^{4–10}. Recent advances in *in situ* techniques now open up the possibility of studying gas–solid interactions at the atomic level^{11–12}. Here we present time-resolved, high-resolution *in situ* transmission electron microscope observations of the formation of carbon nanofibres from methane decomposition over supported nickel nanocrystals. Carbon nanofibres are observed to develop through a reaction-induced reshaping of the nickel nanocrystals. Specifically, the nucleation and growth of graphene layers are found to be assisted by a dynamic formation and restructuring of mono-atomic step edges at the nickel surface. Density-functional theory calculations indicate that the observations are consistent with a growth mechanism involving surface diffusion of carbon and nickel atoms. The finding that metallic step edges act as spatio-temporal dynamic growth sites may be important for understanding other types of catalytic reactions and nanomaterial syntheses.

Here we formed carbon nanofibres by catalytic decomposition of methane over a catalyst consisting of Ni nanoclusters supported on MgAl_2O_4 . This type of catalyst is widely used industrially in the steam reforming process³. The experiments are performed in an *in situ* transmission electron microscope (TEM)¹³. Nickel nanocrystals in the metallic state are formed by reduction of the NiO precursor at 500 °C in about 1 mbar H_2 for roughly 1 h in the TEM. The images show that the Ni nanocrystals have a faceted equilibrium shape (Fig. 1a). Methane is then added to the H_2 atmosphere to obtain a 1:1 mixture at a total pressure of about 2 mbar. TEM images show that graphitic nanofibres form with a Ni nanocluster located at the end and that the alignment of the graphene sheets into multi-layered carbon nanofibre structures couples to a reshaping of the Ni nanocluster. The present study considers Ni nanoclusters with diameters in the range of about 5 to 20 nm. Under these growth conditions, the smaller Ni nanoclusters tend to obtain a highly

elongated shape, and tubular carbon structures form with the graphene sheets aligned parallel to the fibre axis (Fig. 1b). The larger Ni nanoclusters tend to obtain a pear-shape, and graphitic nanofibre structures emerge with the graphene layers inclined with respect to the fibre axis (Fig. 1c).

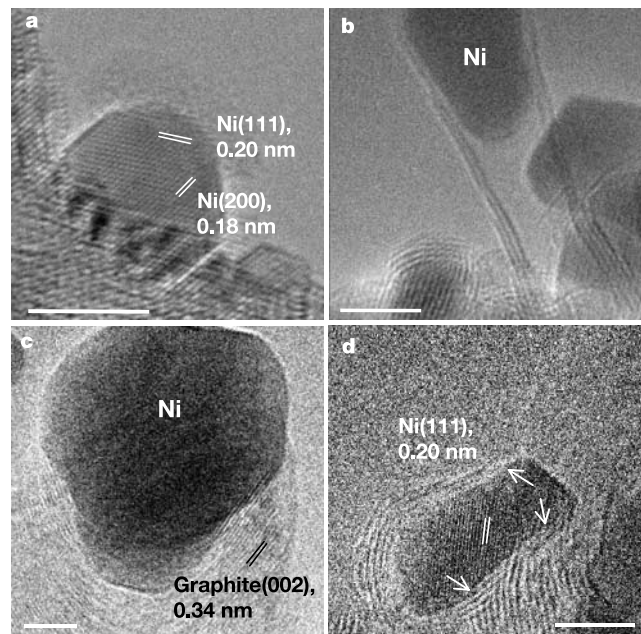


Figure 1 Electron micrographs of the Ni catalyst and carbon nanofibres. **a**, TEM image of a Ni nanocrystal supported on MgAl_2O_4 acquired *in situ* (3.5 mbar H_2 , 430 °C). The lattice fringes correspond to the (111) and (200) lattice planes in metallic Ni. **b**, TEM image showing a multi-walled tubular carbon nanofibre structure. **c**, *In situ* TEM image showing a whisker-type carbon nanofibre. The lattice fringes in the nanofibre correspond to the (002) planes in graphite. **d**, *In situ* TEM image showing a Ni nanocrystal during carbon nanofibre growth. The lattice fringes in the Ni nanocrystal correspond to the (111) planes in metallic Ni. Arrows indicate mono-atomic step edges at the Ni surface. For **b–d**, the reduced Ni catalyst is exposed to a mixture of $\text{CH}_4:\text{H}_2 = 1:1$ at a total pressure of about 2.0 mbar and a temperature of 500–540 °C. Scale bars, 5 nm.

To follow the growth dynamically, a large number of consecutive TEM images of the catalyst have been recorded and visualized as a TEM movie (see Supplementary movies N1 and N2). The main findings in the growth scenario are demonstrated in Fig. 2. The initial equilibrium shape transforms into a highly elongated shape. The elongation of the Ni particle appears to be correlated with the formation of more graphene sheets at the graphene–Ni interface with their basal (002) planes oriented parallel to the Ni surface. Hence, the reshaping of the Ni nanocluster assists the alignment of graphene layers into a tubular structure. The elongation of the Ni nanocrystal continues until it achieves an aspect ratio (length/width) of up to ~ 4 , before it abruptly contracts to a spherical shape within less than ~ 0.5 s (Fig. 2h). The contraction is attributed to the fact that the increase in the Ni surface energy can no longer be compensated for by the energy gained when binding the graphitic fibre to the Ni surface. The elongation/contraction scenario continues in a periodic manner as the nanofibre grows. Incorporation of defects and disordering in the alignment of the graphene sheets can be observed (see Supplementary movie N1). Furthermore, the growth ceases if the graphene layers eventually encapsulate the Ni particle completely, indicating that part of the Ni surface must have direct contact with the gas phase.

During the growth process, lattice-resolved TEM images reveal the detailed structure of the Ni nanoclusters. Such images are typically obtained only in parts of a growth sequence in which the Ni nanocluster is oriented with a zone axis close to the direction of the electron beam (Supplementary movie N2). The image in Fig. 1d demonstrates lattice fringes with a periodicity of 0.20 ± 0.01 nm, corresponding to the (111) planes in Ni. From several sequences of lattice-resolved images, we conclude that the Ni particles are crystalline during the growth scenario.

The high-resolution TEM images furthermore reveal that mono-atomic steps are present at the Ni surface and that a graphene sheet terminates at each of these steps (see Fig. 1d). Such Ni step edges play a key role in the nucleation and growth of graphene sheets: Ni step edges are induced spontaneously in the course of the reaction even at the graphene–Ni interface (Fig. 2b–g; Supplementary movie N2). Between a pair of such step edges, an additional graphene layer grows as the Ni steps move towards the ends of the Ni cluster and vanish. Clearly, this process involves transport of C atoms towards and Ni atoms away from the graphene–Ni interface. The flux of Ni

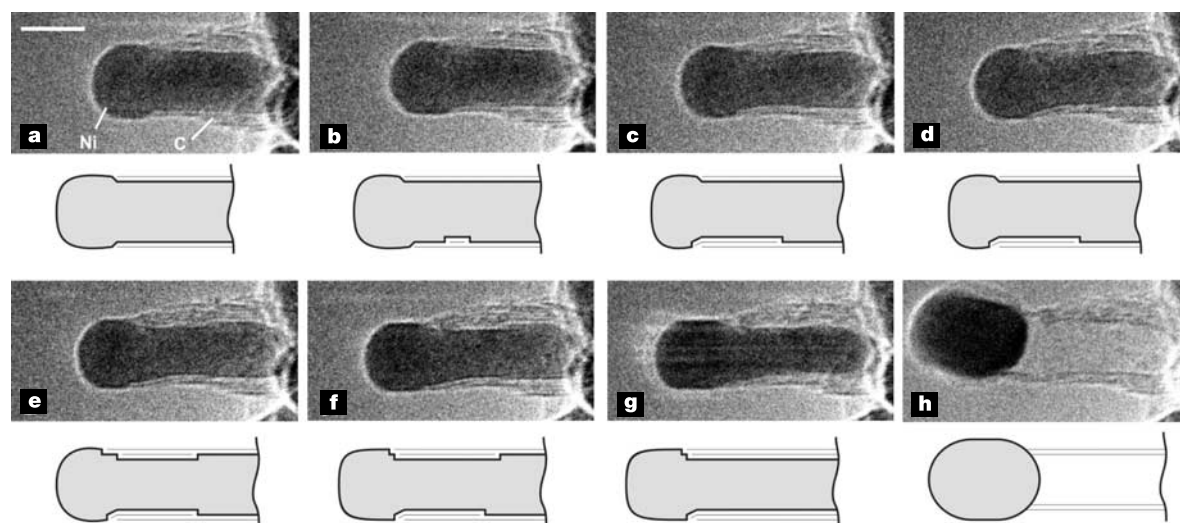


Figure 2 Image sequence of a growing carbon nanofibre extracted from movie N1. Images **a–h** illustrate the elongation/contraction process. Drawings are included to guide the eye in locating the positions of mono-atomic Ni step edges at the C–Ni interface. The

images are acquired *in situ* with $\text{CH}_4:\text{H}_2 = 1:1$ at a total pressure of 2.1 mbar with the sample heated to 536 °C. All images are obtained with a rate of 2 frames s^{-1} . Scale bar, 5 nm.

atoms is directed towards the free Ni surface, producing the observed elongation. The TEM images represent two-dimensional projections of the sample, and thus the observations are indicative of graphene growth around the Ni nanocrystals. The graphene layer appears to extend around the Ni particles in an asymmetric fashion, because interfacial step edge formation can be observed at either side of the Ni nanocrystal at different instants (Fig. 2b, e). This is also supported by the asymmetric shape of the multi-walled tubular nanofibre, which we attribute to an incomplete closure of the outermost graphene sheets.

To understand the origin of the interfacial processes, we have performed density-functional theory (DFT) calculations on the transport of C and Ni adatoms along the graphene–Ni interface. The interface is modelled by a Ni(111) surface with a graphene overlayer. Using the local density approximation (LDA) for exchange and correlation¹⁴, the adhesion energy of the graphene overlayer is calculated to be 0.05 eV per C atom and the binding distance of the overlayer relative to the Ni(111) surface is calculated to be 3.2 Å, in reasonable agreement with experiments¹⁵. We find diffusion energy barriers of 0.1 eV for Ni and 0.5 eV for C adatoms at the graphene–Ni(111) interface (corresponding to paths 1 → 2 and 3 → 4, respectively, in Fig. 3). These values are only slightly different from the diffusion barriers at the free Ni(111) surface (path 5 → 6 for Ni and path 8 → 9 for C). These diffusion energy barriers are found to be essentially the same using the generalized gradient approximation (GGA-RPBE) for exchange and correlation¹⁶.

The main effect of the graphene overlayer is to change the adsorption energy of C and Ni adatoms on the Ni(111) surface. Nickel adatoms are found to bind more strongly by 1.2 eV at the interface than on the free Ni(111) surface (step 1 → 5 in Fig. 3). Previous DFT calculations show that Ni step edges can be induced by adsorbed C atoms because the C binding energy to the Ni step is larger than the energy cost for step formation¹⁷. The graphene overlayer thus helps the formation of Ni steps, and hence, the release of Ni adatoms, which can diffuse along the interface towards the free surface. On the other hand, C adatoms at the interface are destabilized by 0.4 eV (step 8 → 3). This adds to the barrier for carbon diffusion from the free surface to the interface. For the free Ni surface, DFT calculations show that dissociative methane adsorption is facilitated at the step edges and that C atoms adsorb preferentially at the step sites¹⁷. The transport of C atoms from the free Ni surface to sites at the graphene–Ni interface can therefore be

viewed as consisting of the following three steps: the breaking of the C–bond to the Ni–step on the free surface, incorporation under the graphene sheet, and diffusion at the graphene–Ni interface. The energy required to move a C atom from a step to a terrace site on Ni(111) is 0.7 eV (step 7 → 8 in Fig. 3). The extra energy required to incorporate the C atom under the graphene sheet is 0.4 eV (step 8 → 3) and the barrier for C diffusion at the interface is 0.5 eV (path 3 → 4). From this, we estimate the total maximum barrier for incorporation of C atoms to be 1.6 eV. This is in reasonable agreement with the measured activation energy barrier for carbon nanofibre growth in the range 1.3–1.5 eV (refs 8, 18). We therefore suggest that surface transport of C atoms is the rate-limiting step for the nanofibre growth.

It has been suggested that carbon transport could also proceed through the bulk of the Ni particles^{8,18–21}. Although we cannot exclude that such an effect may contribute under some growth conditions, the present results suggest that it is not necessary to include bulk diffusion to obtain a consistent picture. Furthermore, we have established that it is possible to transport C along the graphene–Ni interface. The experiments clearly point to the necessity for transport of Ni. This cannot take place through the bulk metal²², and it is shown that an interface transport mechanism consistent with the C transport mechanism can also explain this phenomenon. According to the DFT calculations, the driving force for graphene formation is related to an energy gain of 0.7 eV per C atom associated with the incorporation of C atoms from the step edges into a graphene sheet adsorbed at Ni(111) (see Fig. 3). In the continuous growth of carbon nanofibres, the graphene layers nucleate and grow preferentially at the graphene–Ni interface. We suggest that the migration of C atoms towards the graphene–Ni interface is driven by an additional energy gain arising from binding a new graphene layer to the existing graphite matrix.

Our direct observations and DFT calculations demonstrate that step edges act as growth centres for graphene growth mainly because carbon binds more strongly to such sites than to sites at the close-packed facets on Ni. The preference of C atoms for step site adsorption is found in general²³, suggesting that the mechanism is of general importance for metal-catalysed nanofibre growth. The finding of a spatiotemporal dynamic growth centre may have further implications for understanding catalytic reactions and nanomaterial syntheses, which often assume a fixed number of stationary active sites. □

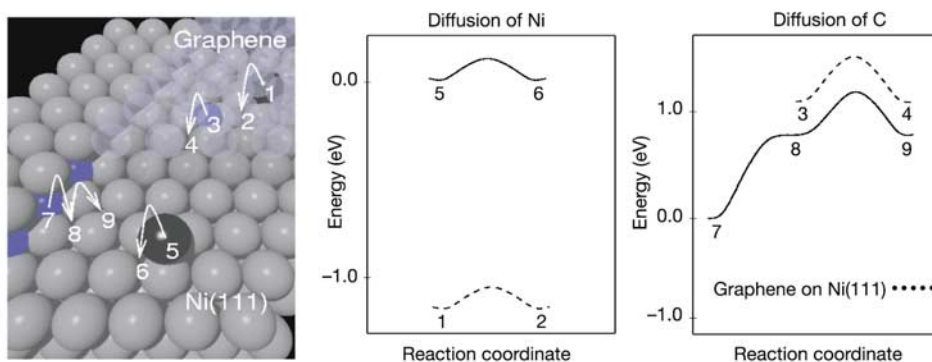


Figure 3 Illustration of the surface processes on the Ni(111) surface modelled in the DFT calculations. The part of the surface with a transparent overlayer indicates the graphene–Ni(111) interface. Diffusion of a Ni adatom (black) between threefold hollow sites is indicated at the interface, path (1 → 2), and at the uncovered Ni(111) surface, path (5 → 6). Diffusion of a C adatom (blue) is indicated between threefold hollow sites at the

interface, path (3 → 4), and from the step edge to threefold hollow sites at the uncovered Ni(111) terrace, path (7 → 8 → 9). The calculated energies along the depicted diffusion paths are shown for a Ni and C adatom at the Ni(111) surface (solid curves) and at the graphene–Ni(111) interface (dashed curves). The energy of a C atom in a graphene sheet on Ni(111) is shown by the dotted line.

Methods

The growth experiments were carried out in a Philips CM300 FEG TEM equipped with an *in situ* cell¹³. This system provides images with a resolution better than 0.14 nm of samples exposed to reactive gases and elevated temperatures. A Tietz FastScan-F114 CCD was used for recording TEM images. The image acquisition time was adjusted to suit the reshaping of the Ni nanocrystals and ranged from 2 to 10 frames s⁻¹. A total of 16 video sequences (each consisting of 300–1,200 consecutive high-resolution images) showing the coupling of graphene growth to Ni step-edge dynamics were recorded in six different growth experiments. The electron beam intensity was kept in the interval 0.2–2 Å cm⁻², which was sufficiently low to avoid any influence of the electron beam on the growth. Growth experiments with the electron beam switched off resulted in the same carbon nanofibre structures as in the low-dose *in situ* experiments.

The self-consistent DFT calculations were performed using the LDA¹⁴ or the GGA-RPBE approximation¹⁶ for exchange and correlation. The interaction of the graphene overlayer with the Ni(111) surface was dominated by van der Waals forces. The LDA for the exchange and correlation term is known to give the best description of such weak interactions²⁴. We have therefore used LDA for calculating the adsorption energy and diffusion energy barrier for C and Ni adatoms on the clean Ni(111) surface and the graphene–Ni(111) interface. The adsorption energies and diffusion energy barriers on the clean Ni(111) surface calculated by the GGA-RPBE approximation were found to be essentially the same. However, within the GGA-RPBE approximation we found no adhesion of the graphene overlayer to the Ni(111) surface.

The ionic cores are described by ultrasoft pseudopotentials²⁵ and the one-electron valence states were expanded in a basis of plane waves with kinetic energy below 25 Ry. The electron density was determined self-consistently by iterative diagonalization of the Kohn–Sham hamiltonian, Pulay mixing²⁶ of the resulting electronic density and Fermi occupation of the Kohn–Sham states ($k_B T = 0.1$ eV). All total energy calculations were extrapolated to zero electronic temperature and performed with the magnetic moment of the nickel surface taken into account.

The Ni(111) surface was modelled by a periodic repetition of a three-layer slab, whereas a slab thickness of nine layers was used in the [211] direction to create a stepped surface and a thickness comparable to the Ni(111) surface. In both cases, the slabs were separated by ~ 10 Å of vacuum. In the Ni(111) model, the top layer was allowed to fully relax, whereas in the Ni(211) model, the uppermost close-packed (111) layer was relaxed fully. The remaining layers were kept fixed in the bulk geometry with a calculated equilibrium lattice constant of 3.52 Å. For all calculations on Ni(111), a $p(2 \times 2)$ lateral unit cell was used, giving C and Ni adsorbate coverages of 0.25 ML, whereas the C/Ni ratio for the graphene overlayer was 2. For the Ni(211) surface, a (1×2) unit cell was used in the calculations with a one-to-one coverage of adsorbate to Ni step edge atoms. A sampling of $4 \times 4 \times 1$ Monkhorst–Pack special k -points was used, resulting in 8 k -points in the irreducible Brillouin zone.

Received 20 August; accepted 5 December 2003; doi:10.1038/nature02278.

- Baughman, R. H., Zakhidov, A. A. & de Heer, W. A. Carbon nanotubes—the route towards applications. *Science* **297**, 787–792 (2002).
- Ajayan, P. M. Nanotubes from carbon. *Chem. Rev.* **99**, 1787–1799 (1999).
- Rostrup-Nielsen, J. R., Sehested, J. & Nørskov, J. K. Hydrogen and synthesis gas by steam and CO₂ reforming. *Adv. Catal.* **47**, 65–138 (2002).
- Ebbesen, T. W. & Ajayan, P. M. Large-scale synthesis of carbon nanotubes. *Nature* **358**, 220–222 (1992).
- Thess, A. *et al.* Crystalline ropes of metallic carbon nanotubes. *Science* **273**, 483–487 (1996).
- Kong, J., Soh, H. T., Cassell, A. M., Quate, C. F. & Dai, H. Synthesis of individual single-walled carbon nanotubes on patterned silicon wafers. *Nature* **395**, 878–881 (1998).
- Boskovic, B. O. *et al.* Large-area synthesis of carbon nanofibres at room temperature. *Nature Mater.* **1**, 165–168 (2002).
- Baker, R. T. K., Harris, P. S., Thomas, R. B. & Waite, R. J. Nucleation and growth of carbon deposits from the nickel catalyzed decomposition of acetylene. *J. Catal.* **26**, 51–62 (1972).
- Rostrup-Nielsen, J. R. Equilibria of decomposition reactions of carbon monoxide and methane over nickel catalysts. *J. Catal.* **27**, 343–356 (1972).
- Charlier, J.-C. & Iijima, S. in *Growth Mechanisms of Carbon Nanotubes* (eds Dresselhaus, M. S., Dresselhaus, G. & Avouris, Ph.) Vol. 80, *Topics in Applied Physics* 55–80 (Springer, Heidelberg, 2001).
- Clausen, B. S., Topsøe, H. & Frahm, R. Application of combined X-ray diffraction and absorption techniques for *in situ* catalyst characterisation. *Adv. Catal.* **42**, 315–344 (1998).
- Topsøe, H. *In situ* characterisation of catalysts. *Stud. Surf. Sci. Catal.* **130**, 1–21 (2000).
- Hansen, P. L. *et al.* Atom-resolved imaging of dynamic shape changes in supported copper nanocrystals. *Science* **295**, 2053–2055 (2002).
- Jones, R. O. & Gunnarsson, O. The density functional formalism, its applications and prospects. *Rev. Mod. Phys.* **61**, 689–746 (1989).
- Rosei, R. *et al.* Structure of graphitic carbon on Ni(111): A surface extended-energy-loss fine-structure study. *Phys. Rev. B* **28**, 1161–1164 (1983).
- Hammer, B., Hansen, L. B. & Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew–Burke–Ernzerhof functionals. *Phys. Rev. B* **59**, 7413–7421 (1999).
- Bengard, H. S. *et al.* Steam reforming and graphite formation in Ni catalysts. *J. Catal.* **209**, 365–384 (2002).
- Trimm, D. L. The formation and removal of coke from nickel catalyst. *Catal. Rev. Sci. Eng.* **16**, 155–189 (1977).
- Alstrup, I. A new model explaining carbon filament growth on nickel, iron and Ni–Cu alloy catalysts. *J. Catal.* **109**, 241–251 (1988).
- Rostrup-Nielsen, J. R. & Trimm, D. L. Mechanisms of carbon formation on nickel-containing catalysts. *J. Catal.* **48**, 155–165 (1977).
- De Jong, K. P. & Geus, J. W. Carbon nanofibers: catalytic synthesis and applications. *Catal. Rev. Sci. Eng.* **42**, 481–510 (2000).
- Askeland, D. R. *The Science and Engineering of Materials* 2nd edn, Ch. 5 (PWS-KENT, Boston, 1989).

- Nørskov, J. K. *et al.* Universality in heterogeneous catalysis. *J. Catal.* **209**, 275–278 (2002).
- Da Silva, J. L. F., Stampfl, C. & Scheffler, M. Adsorption of Xe atoms on metal surfaces: new insights from first-principles calculations. *Phys. Rev. Lett.* **90**, 066104 (2003).
- Vanderbilt, D. H. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **41**, 7892–7895 (1990).
- Kresse, G. & Furthmüller, J. Efficiency of *ab-initio* total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We gratefully acknowledge the participation of the CTCI Foundation, Taiwan, in the establishment of the TEM facility at Haldor Topsøe A/S. S.H. acknowledges support from the Danish Research Council, STVF. F.A.P. and J.K.N. acknowledge computational resources from the Danish Center for Scientific Computing. C.L.-C. acknowledges support from Facultad de Ciencias, Universidad de Cádiz, Puerto Real, Spain, for a sabbatical visit to Haldor Topsøe A/S.

Competing interests statement The authors declare competing financial interests: details accompany the paper on www.nature.com/nature.

Correspondence and requests for materials should be addressed to S.H. (sth@topsoe.dk).

Enhanced ice sheet growth in Eurasia owing to adjacent ice-dammed lakes

G. Krinner¹, J. Mangerud², M. Jakobsson³, M. Crucifix^{4*}, C. Ritz¹ & J. I. Svendsen²

¹LGGE, CNRS-UJF Grenoble, BP 96, F-38402 Saint Martin d'Hères, France

²Department of Earth Science, University of Bergen, and Bjerknes Centre for Climate Research, Allég. 41, N-5007 Bergen, Norway

³Center for Coastal and Ocean Mapping/Joint Hydrographic Center, Chase Ocean Engineering Laboratory, 24 Colovos Road, Durham, New Hampshire 03824, USA

⁴Institut d'Astronomie et de Géophysique G. Lemaître, Université catholique de Louvain, chemin du Cyclotron, 2, B-1348 Louvain-la-Neuve, Belgium

* Present address: Met Office, Hadley Centre for Climate Prediction and Research, Fitzroy Road, Exeter EX1 3PB, UK

Large proglacial lakes cool regional summer climate because of their large heat capacity, and have been shown to modify precipitation through mesoscale atmospheric feedbacks, as in the case of Lake Agassiz¹. Several large ice-dammed lakes, with a combined area twice that of the Caspian Sea, were formed in northern Eurasia about 90,000 years ago, during the last glacial period when an ice sheet centred over the Barents and Kara seas² blocked the large northbound Russian rivers³. Here we present high-resolution simulations with an atmospheric general circulation model that explicitly simulates the surface mass balance of the ice sheet. We show that the main influence of the Eurasian proglacial lakes was a significant reduction of ice sheet melting at the southern margin of the Barents–Kara ice sheet through strong regional summer cooling over large parts of Russia. In our simulations, the summer melt reduction clearly outweighs lake-induced decreases in moisture and hence snowfall, such as has been reported earlier for Lake Agassiz¹. We conclude that the summer cooling mechanism from proglacial lakes accelerated ice sheet growth and delayed ice sheet decay in Eurasia and probably also in North America.

We used the LMDZ3 (Laboratoire de Météorologie Dynamique, CNRS Paris) stretched-grid atmosphere general circulation model (AGCM). In addition to a present-day reference simulation (called 'PD'), two simulations for 90 kyr before present (BP) were carried out. Both were constrained by appropriate ice sheet reconstructions, sea surface conditions, greenhouse gas concentrations and insolation as boundary conditions (Fig. 1; see Methods for details). They

differ by the fact that experiment 'L' (for 'lakes') includes Russian ice-dammed lakes³ and the Baltic lake⁴, while these are replaced by normal continental surface in simulation 'NL' (for 'no lakes'). The grid-stretching capability of LMDZ3 allows high resolution (100 km) over the Barents–Kara region. Compared with observed present-day surface air temperature⁵ and precipitation⁶, the mean absolute annual temperature error in simulation PD over the region of interest is fairly low (1.0 °C), but annual precipitation is over-estimated by about 60%, mainly because of a wet bias in winter. On the other hand, LMDZ3 quite adequately captures the present surface mass balance patterns and quantities over the present ice sheets^{7,8}.

Analysis of the upper soil temperatures in the region in simulation L, based on a normalized surface frost index⁹, shows that the Russian proglacial lakes were mostly situated within the limit of continuous permafrost, only the southernmost part of the West Siberian Lake (Fig. 1) extending into a region with extensive discontinuous permafrost. This is in agreement with ice wedge casts found on palaeo-beaches¹⁰. In both L and NL, simulated precipitation 90 kyr ago is less than 50% of PD over large parts of Russia. Simulated inflow of ice sheet melt water is substantial in L, corresponding to more than 1 myr⁻¹ lake level rise. This is supported by findings of ice marginal deltas formed in the lakes¹¹. Simulated lake summer surface temperatures exceed 4 °C only near the southern shore. In their northern parts, the proglacial lakes are cold monomictic (one vertical mixing per year, in summer; bottom temperature below 4 °C). They remain ice-covered from 7 months per year on the southern shore to 11 months per year near the ice sheet in spite of high meltwater input rates. The Baltic Lake, situated at lower latitudes, is somewhat warmer, and the ice cover lasts between 6 months on the southern shore and 9 months near the ice sheet. During all seasons, surface winds over the ice sheet in L and NL show a typical katabatic pattern, the air masses flowing down the slope of the ice sheet, deviated to the right because of the Earth's rotation, with maximum speeds near the ice sheet margin.

The primary atmospheric effect of the lakes is a strong summer cooling due to their large heat capacity. Summer surface air temperature anomalies (L–NL) exceed –10 °C above the lakes. The impact over the southern margin of the Barents–Kara ice sheet is substantial up to an altitude of about 2,000 m (Fig. 2a). This cooling effect extends horizontally about 1,000 km away from the lakes. The accumulation of cool air off and above the ice sheet

margin decreases the horizontal air density gradient, and therefore induces a weakening of the katabatic circulation in summer. Above the lakes, the summer cooling is associated with a substantial surface pressure increase (+4 mbar). The summer cooling, and, to a lesser degree, the reduced surface wind speed on the ice sheet slope (leading to lower downward turbulent surface heat fluxes), strongly reduce summer melt along the southern margin of the Barents–Kara ice sheet and the southeastern margin of the Scandinavian ice sheet. On the lakeward parts of the ice sheet, runoff is reduced by 8 mm d⁻¹ in July, and by more than 500 mm yr⁻¹ (50%) in the annual mean, when the lakes are taken into account (refreezing of melt water is accounted for in calculating runoff from snow and ice melt¹²). Sensitivity tests showed that both the sign and the absolute value of the annual mean precipitation change are dependent on the prescribed sea-surface boundary conditions, but the precipitation change always remains far smaller than the ablation decrease. The latter coherently dominates the simulated surface mass balance change (Fig. 2b).

The Barents–Kara and the Scandinavian ice sheets were linked by a relatively small ice bridge at about 35 °E. Their surface mass balances (SMB, Fig. 3) are therefore analysed separately. The integrated SMB of the Barents–Kara ice sheet, defined as precipitation minus evaporation/sublimation and minus runoff, is +0.18 m of water equivalent per year (m_w yr⁻¹) in L, and –0.43 m_w yr⁻¹ in NL. According to a paired Student's *t*-test with nine degrees of freedom, the confidence in the SMB increase in L with respect to NL exceeds 99%, which excludes our result from being an artefact of interannual variability. Furthermore, the southern ablation zone of the Barents–Kara ice sheet extends over

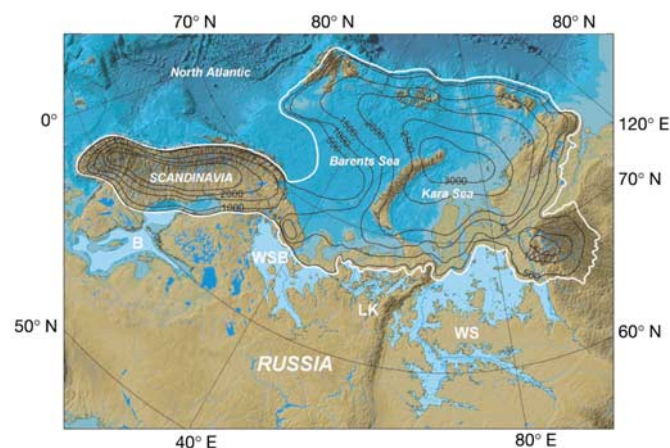


Figure 1 Maximum ice sheet and lake extent in northern Russia during the early Weichselian, about 90,000 yr ago. Abbreviations denote the West Siberian Lake (WS), Lake Komi (LK), the White Sea Basin Lake (WSB), and the Baltic Lake (B).

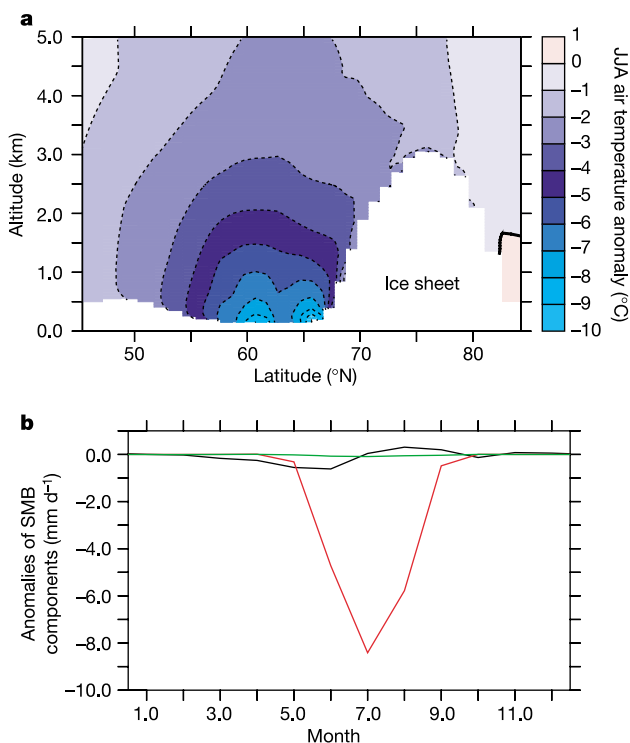


Figure 2 Lake-induced climate anomalies between simulation L and simulation NL. **a**, JJA (June, July, August) air temperature anomalies: vertical cut through the lower atmosphere at 70° E. The pattern is representative for all longitudes where proglacial lakes are adjacent to the Barents–Kara ice sheet. **b**, Monthly mean anomalies of surface mass balance (SMB) components precipitation (black), runoff (red) and evaporation/sublimation (green) on the Scandinavian/Barents–Kara ice sheet grid points that belong to the lakes' drainage basin.

several grid points towards the interior in simulation NL, and significant melt (in excess of $0.2 \text{ m}_w \text{ yr}^{-1}$) occurs in both simulations at least two grid points from the ice limit. Hence, the resolution is sufficient to capture melt processes, such that the sensitivity of the SMB can be better trusted than in most AGCMs that hardly represent the ablation zone. Concerning the Scandinavian ice sheet, the Russian proglacial lakes and the Baltic Lake induce a substantial increase ($+1.7 \text{ m}_w \text{ yr}^{-1}$) of the simulated SMB owing to reduced melt rates. Note that the prescribed western margin of the Scandinavian ice sheet does not reach the Atlantic coastline, and that outlet glaciers in fjords and valleys are not prescribed, owing to insufficient resolution. The westernmost ice sheet grid points are therefore at altitudes (more than 1,000 m, mostly more than 1,500 m) where significant ablation does not occur, which explains the high surface mass balance obtained here. High melt rates and therefore negative surface mass balances are simulated on a few low-altitude grid points near Franz Josef Land. Given the high latitude, the great distance to the Atlantic, the vicinity of the glacial Arctic Ocean, the ice sheet to the south, and low CO_2 concentrations, this result seems unlikely.

Although it focused on a different climatic and regional context, it is interesting to note that a study of the North American climate at 11,000 yr BP (ref. 1) suggested that the presence of proglacial Lake Agassiz caused a reduction of precipitation rates on the adjacent Laurentide ice sheet (the magnitude of which is similar to the precipitation changes induced by the Russian proglacial lakes, and much weaker than the ablation reduction reported here). The authors concluded that this precipitation reduction led to a

decreased surface mass balance of the Laurentide ice sheet. However, assessing the surface mass balance precisely was not possible in that study as ice melt was not explicitly calculated (S. W. Hostetler, personal communication). Because the same study suggested that Lake Agassiz induced a significant regional cooling in summer, similar to the cooling induced by the Russian lakes reported here, we postulate that the lake-induced reduction in melting on the southern margin of the Laurentide ice sheet was larger than the reduction of precipitation. Furthermore, we see no reason why the feed-back mechanism reported here should not have operated during earlier growth and decay phases of the Laurentide ice sheet, during which large ice-dammed lakes are known to have existed¹³.

The main conclusion of this study is that the Russian proglacial lakes, through their effect on the surface mass balance of the Barents–Kara ice sheet, repeatedly played an important role in the regional climate dynamics. Inception of the Barents–Kara ice sheet was probably caused by insolation changes and positive feedbacks involving snow albedo, ocean circulation¹⁴, vegetation^{15,16} and greenhouse gas concentration changes¹⁷. As soon as the ice sheet was big enough to dam the northward flowing Russian rivers, thereby creating the proglacial lakes, this initial growth of the Barents–Kara ice sheet was notably enhanced. Optical stimulated luminescence dates indicate that the ice sheet attained its maximum extent between 90 and 80 kyr ago^{2,10}. On the other hand, June insolation at 65°N , which is close to the southern limit of the Barents–Kara ice sheet, increased by $\sim 12\%$ between 95 and 85 kyr ago¹⁸. The most coherent scenario is therefore that the ice sheet attained its maximum extent at about 90 kyr ago, and that a subsequent retreat (delayed, but not stopped by the lake-induced melt reduction) began around 85 kyr ago owing to increased summer insolation. Once the Barents–Kara ice sheet had shrunk sufficiently, rapid subglacial drainage of the lakes occurred. This further accelerated the deglaciation by leading to higher summer melt rates along the southern ice sheet margin.

Note that rapid drainage of proglacial lakes along the southern margin of the Laurentide ice sheet is thought to have caused abrupt widespread cooling by interrupting the oceanic thermohaline circulation¹⁹; it has been suggested that a similar cooling, possibly compensating for the regional warming caused by the drainage of the lakes, might have occurred following lake drainage in northern Eurasia¹¹. During later stages of the last glacial, in particular the Last Glacial Maximum (about 20,000 yr ago), the Barents–Kara ice sheet did not grow big enough to create large proglacial lakes¹¹ (possibly because the larger Scandinavian ice sheet blocked moisture delivery from the Atlantic), and the growth acceleration due to these lakes therefore did not take place. Finally, we note that the coupled lake–ice sheet system could be self-sustaining to a certain degree: a small ice sheet decay leads to an increased lake size, thereby cooling the climate and thus reducing the meltwater production. This negative feedback will break down if the ice sheet decay is sufficiently pronounced to allow drainage of the lake through previously blocked passages. □

Methods

Model

The LMDZ3 model has been run with 96×72 (longitude $\times$ latitude) grid points. The chosen nominal horizontal resolution is irregular, varying from 100 km at the centre of the region of interest (between 0° and 120°E and between 55° and 85°N) to about 550 km roughly at the antipodes of this region. The model has 19 vertical layers in the atmosphere. LMDZ3 has been adapted for the simulation of cold climates²⁰ and includes a thermal lake module²¹.

Boundary conditions

The prescribed boundary conditions (Fig. 1) for the AGCM include orbital parameters for 90 kyr ago¹⁸, an atmospheric CO_2 concentration of 205 p.p.m.v. (ref. 22); reconstructed dimensions of the Barents–Kara and the Scandinavian ice sheet based on geological evidence² and glaciological modelling²³ with some modifications to fit the latest results

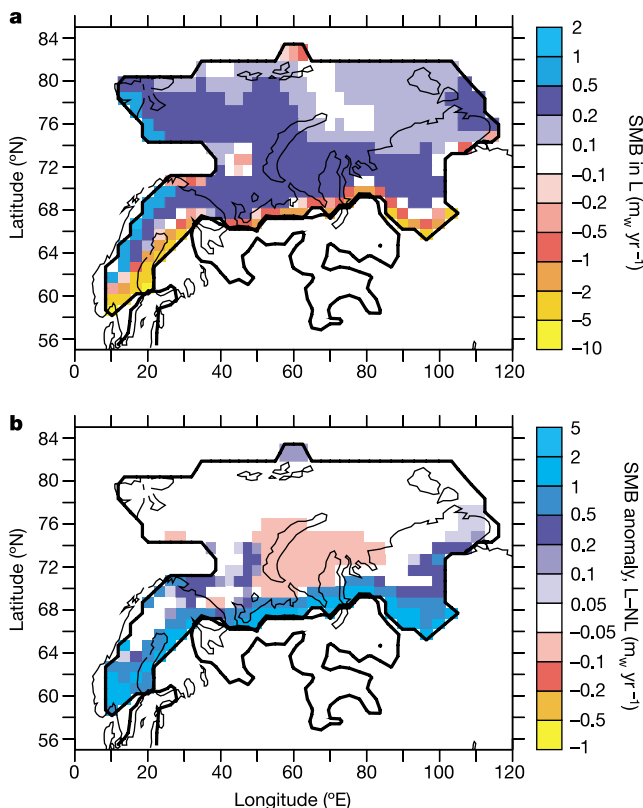


Figure 3 Simulated ice sheet surface mass balance and lake-induced surface mass balance anomaly. **a**, Simulated surface mass balance in simulation L; **b**, lake-induced surface mass balance anomaly (L–NL). Units are metres of water equivalent per year. Bold lines, ice sheets and lakes (grid-scale fraction $> 20\%$) for 90 kyr ago; thin line, present-day coastline.

obtained within the ESF-QUEEN project; ice sheet thicknesses for the rest of the Northern Hemisphere obtained using the climatic index method²⁴ applied to an ice sheet model and AGCM outputs²⁵; modelled global sea surface temperatures and sea-ice extent²⁶; surface albedo and roughness over ice-free surfaces derived from modelled glacial maximum palaeovegetation²⁷; a large lake in the part of the Baltic Sea depression not covered by the Scandinavian ice sheet²⁸; and reconstructed proglacial lakes in European Russia and in West Siberia³.

Initial conditions and spin-up

The simulations were started from an initial state with present-day soil temperatures. From the 15 yr of each simulation, the first five years were discarded as spin-up. Lake temperatures at the bottom level (between 15 and 30 m) were in equilibrium at the end of the 5 spin-up years. At the end of each of the first three spin-up years for the 90-kyr-ago simulations, soil temperature, being the slowest simulated component of the climate system simulated here, was separately spun up for 1,000 yr using the year's modelled monthly surface temperatures.

Received 15 May; accepted 14 November 2003; doi:10.1038/nature02233.

- Hostetler, S. W., Bartlein, P. J., Clark, P. U., Small, E. E. & Solomon, A. M. Simulated influences of Lake Agassiz on the climate of central North America 11,000 years ago. *Nature* **405**, 334–337 (2000).
- Svendsen, J. I. *et al.* Late Quaternary ice sheet history of Eurasia. *Quat. Sci. Rev.* (in the press).
- Mangerud, J., Astakhov, V., Jakobsson, M. & Svendsen, J. I. Huge ice-age lakes in Russia. *J. Quat. Sci.* **16**, 773–777 (2001).
- Lundquist, J. Glacial stratigraphy in Sweden. *Spec. Pap. Geol. Surv. Finl.* **15**, 43–59 (1992).
- Legates, D. R. & Willmott, C. J. Mean seasonal and spatial variability in global surface air temperature. *Theor. Appl. Climatol.* **41**, 11–21 (1990).
- Legates, D. R. & Willmott, C. J. Mean seasonal and spatial variability in gauge-corrected, global precipitation. *Int. J. Climatol.* **10**, 111–127 (1990).
- Genthon, C. & Krinner, G. Antarctic surface mass balance and systematic biases in general circulation models. *J. Geophys. Res.* **106**, 20653–20664 (2001).
- Krinner, G. & Werner, M. Impact of precipitation seasonality changes on isotopic signals in polar ice cores: A multi-model analysis. *Earth Planet. Sci. Lett.* **216**, 525–538 (2003).
- Nelson, F. & Outcalt, S. I. A computational method for prediction and regionalization of permafrost. *Arct. Alp. Res.* **19**, 279–288 (1987).
- Mangerud, J., Astakhov, V., Murray, A. & Svendsen, J. I. The chronology of a large ice-dammed lake and the Barents–Kara ice sheet advances, Northern Russia. *Glob. Planet. Change* **31**, 321–336 (2001).
- Mangerud, J. *et al.* Ice-dammed lakes and rerouting of the drainage of northern Eurasia during the last glaciation. *Quat. Sci. Rev.* (in the press).
- Thompson, S. L. & Pollard, D. Greenland and Antarctic mass balances for present and doubled atmospheric CO₂ from the GENESIS version-2 global climate model. *J. Clim.* **10**, 871–900 (1997).
- Teller, J. T. in *The Quaternary Period in the United States* (eds Gillespie, A., Porter, S. & Atwater, B.) Ch. 3 (Elsevier, Amsterdam, 2003).
- Khodri, M. *et al.* Simulating the amplification of orbital forcing by ocean feedbacks in the last glaciation. *Nature* **410**, 570–574 (2001).
- de Noblet, N. *et al.* Possible role of atmosphere-biosphere interactions in triggering the last glaciation. *Geophys. Res. Lett.* **23**, 3191–3194 (1996).
- Gallimore, R. G. & Kutzbach, J. E. Role of orbitally induced changes in tundra area in the onset of glaciation. *Nature* **381**, 503–505 (1996).
- Shackleton, N. J. The 100,000-year ice-age cycle identified and found to lag temperature, carbon dioxide, and orbital eccentricity. *Science* **289**, 1897–1902 (2000).
- Berger, A. Long-term variations of daily insolation and Quaternary climatic changes. *J. Atmos. Sci.* **35**, 2362–2367 (1978).
- Clarke, G., Leverington, D., Teller, J. & Dyke, A. Superlakes, megafloods, and abrupt climate change. *Science* **301**, 922–923 (2003).
- Krinner, G., Genthon, C., Li, L. & Le Van, P. Studies of the Antarctic climate using a stretched-grid general circulation model. *J. Geophys. Res.* **102**, 13731–13745 (1997).
- Krinner, G. Impact of lakes and wetlands on boreal climate. *J. Geophys. Res.* **D 108**, 101029/2002JD002597 (2003).
- Petit, J.-R. *et al.* Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**, 426–436 (1999).
- Siebert, M. J., Dowdeswell, J. A., Hald, M. & Svendsen, J. I. Modelling the Eurasian ice sheet through a full (Weichselian) glacial cycle. *Glob. Planet. Change* **31**, 367–385 (2001).
- Marshall, S. J., Tarasov, L., Clarke, G. & Peltier, W. R. Glaciological reconstruction of the Laurentide Ice Sheet: Physical processes and modelling challenges. *Can. J. Earth Sci.* **37**, 769–793 (2000).
- Charbit, S., Ritz, C. & Ramstein, G. Simulations of Northern Hemisphere ice-sheet retreat: Sensitivity to physical mechanisms involved during the Last Deglaciation. *Quat. Sci. Rev.* **21**, 243–266 (2002).
- Crucifix, M., Loutre, M. F., Tulkens, P., Fichefet, T. & Berger, A. Climate evolution during the Holocene: a study with an Earth system model of intermediate complexity. *Clim. Dyn.* **19**, 43–60 (2002).
- Crowley, T. Ice age terrestrial carbon changes revisited. *Glob. Biogeochem. Cycles* **9**, 377–389 (1995).

Acknowledgements We thank S. Hostetler for discussions and M. Siebert for comments and suggestions. Model simulations were carried out at IDRIS/CNRS. This work was supported by the ESF and the French national programmes ECLIPSE, PNEDC and ACI Jeunes Chercheurs. The field work and other analyses were funded by the Research Council of Norway by grants to the PECHORA project. M.J. was supported by NOAA.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to G.K. (krinner@ujf-grenoble.fr).

Natural examples of olivine lattice preferred orientation patterns with a flow-normal *a*-axis maximum

Tomoyuki Mizukami^{1,2}, Simon R. Wallis² & Junji Yamamoto^{3*}

¹Department of Geology and Mineralogy, University of Kyoto, Sakyo-ku, Kyoto, 606-8501, Japan

²Department of Earth and Planetary Sciences, Graduate School of Environmental Studies, University of Nagoya, Chikusa-ku, Nagoya, 464-8601, Japan

³Laboratory of Earthquake Chemistry, University of Tokyo, Bunkyo-ku, Tokyo, 113-0033, Japan

* Present address: Department of Earth and Planetary Sciences, Tokyo Institute of Technology, Meguro-ku, Tokyo, 152-8551, Japan

Tectonic plate motion is thought to cause solid-state plastic flow within the underlying upper mantle and accordingly lead to the development of a lattice preferred orientation of the constituent olivine crystals^{1–3}. The mechanical anisotropy that results from such preferred orientation typically produces a direction of maximum seismic wave velocity parallel to the plate motion direction^{4,5}. This has been explained by the existence of an olivine preferred orientation with an '*a*-axis' maximum parallel to the induced mantle flow direction^{3,5,6–8}. In subduction zones, however, the olivine *a* axes have been inferred to be arranged roughly perpendicular to plate motion^{9–13}, which has usually been ascribed to localized complex mantle flow patterns^{10–13}. Recent experimental work¹⁴ suggests an alternative explanation: under conditions of high water activity, a 'B-type' olivine preferred orientation may form, with the *a*-axis maximum perpendicular to the flow direction. Natural examples of such B-type preferred orientation are, however, almost entirely unknown. Here we document widespread B-type olivine preferred orientation patterns from a subduction-type metamorphic belt in southwest Japan and show that these patterns developed in the presence of water. Our discovery implies that mantle flow above subduction zones may be much simpler than has generally been thought.

The Higashi-akaishi peridotite body of southwest Japan is the only kilometre-scale garnet peridotite body yet found in a subduction-type metamorphic belt¹⁵. This body offers an unrivalled opportunity to study subduction zone mantle directly. Here we focus on the olivine lattice preferred orientation (LPO) patterns of this body. The Higashi-akaishi body is tabular with an area of 10 km² and a thickness of 500 m (ref. 16) and occurs as an integral part of the Cretaceous Sanbagawa high-pressure, low-temperature metamorphic belt^{17–19}. Dunite, a rock type of more than 90% olivine, is the main constituent of the body with minor amounts of wehrlite, pyroxenite, garnet pyroxenite, chromitite and garnet peridotite^{15,16,20}. Serpentinized equivalents of these rock types developed in the periphery of the body¹⁴ during exhumation, but the earlier olivine-rich microstructures and associated petrological information are well preserved in the inner part (Figs 1, 2).

Field observations and microstructural analyses allow two generations of tectonic fabrics to be distinguished in the dunite. The earliest D₁ fabric is defined by the crystal-shape preferred orientation of coarse clear olivine grains (~0.6 mm) (Fig. 1a). A second D₂ fabric is widespread throughout the Higashi-akaishi body and consists of coarse dusty olivine porphyroclasts (~0.5 mm) and fine clear olivine neoblasts (~0.1 mm) (Fig. 1c), which formed by dynamic recrystallization of the D₁ coarse-grained fabric (Fig. 1d). The preferred alignment of olivine crystals defines planar and linear fabrics for both D₁ (Fig. 1a) and D₂ (Fig. 1c). The olivine grain shape lineations are interpreted to be parallel to the maximum extension direction. This interpretation is supported by the com-

mon presence of elongate and boudinaged spinel grains with the same orientation. In many cases, field measurements of both planar and linear fabrics were cross-checked by multiple sectioning and microscopic observations of oriented samples. No significant difference was observed between the orientations determined in the field and those determined in the laboratory. The presence of composite planar and linear tectonic fabrics (L–S tectonites) implies roughly plane-strain conditions and, therefore, that the stretching lineation will approximate to the flow direction.

Both D_1 and D_2 are associated with well-developed and distinct olivine LPO patterns. Coarse-grained olivine of the D_1 fabric exhibits a girdle-type LPO with the b -axis concentration normal to the D_1 foliation and with a weak concentration of the a axis subparallel to the D_1 mineral lineation. In contrast, the olivine LPO associated with D_2 shows strong concentrations of the three crystallographic axes: with the c [001] axis subparallel to the stretching lineation, the b [010] axis subperpendicular to the tectonic foliation and the a [100] axis within the foliation and subperpendicular to the stretching lineation (Fig. 2b). This configuration is identical to the B-type LPO of ref. 14. The B-type LPO patterns and characteristic D_2 porphyroclastic microstructure are widely developed throughout the Higashi-akaishi body (Fig. 2a). In total, 12 such B-type fabrics were identified from 11 localities. The strongest B-type LPO patterns are found in the samples with the lowest proportion of D_1 porphyroclastic olivine, which we interpret to be the most deformed. Although there is some difference in strength, the topology of D_2 olivine LPO patterns is similar in both strongly

and weakly deformed samples. We conclude that the LPO patterns, particularly those in the strongly deformed samples, closely approximate to steady-state fabrics and that the effects of earlier D_1 structures are negligible.

Porphyroclastic microstructures are commonly observed in deformed mantle xenoliths and peridotite bodies from various tectonic settings^{21,22}. However, the D_2 porphyroclastic microstructure of the Higashi-akaishi body has several distinct features suggesting that it developed in the presence of water. The olivine porphyroclasts have a dusty appearance owing to micro-inclusions, whereas the neoblasts are clear (Fig. 1d). Similar microstructures have also been described in a previous study²³. There are only very few inclusions in the pre-existing D_1 olivine (Fig. 1c), implying that the formation of micro-inclusions is related to the main D_2 deformation. High-powered microscopic study of 100- μ m-thick sections revealed that most of these inclusions are two-phase inclusions with semi-faceted shapes and are commonly associated with radially arranged cracks in the host olivine crystal (Fig. 3a). Most cracks terminate within the host crystal and do not intersect grain boundaries. Micro-Raman spectroscopic analysis of the micro-inclusions detects a sharp band corresponding to the O–H stretching vibration at $3,650\text{ cm}^{-1}$ in Raman shift (Fig. 3b) with other peaks equivalent to those of serpentine. The sharp high frequency of the O–H stretching is consistent with a very weak hydrogen bond such as the hydroxyl ion of serpentine and not molecular H_2O . These observations suggest that the inclusions are the remains of water-rich fluid inclusions. The formation of the

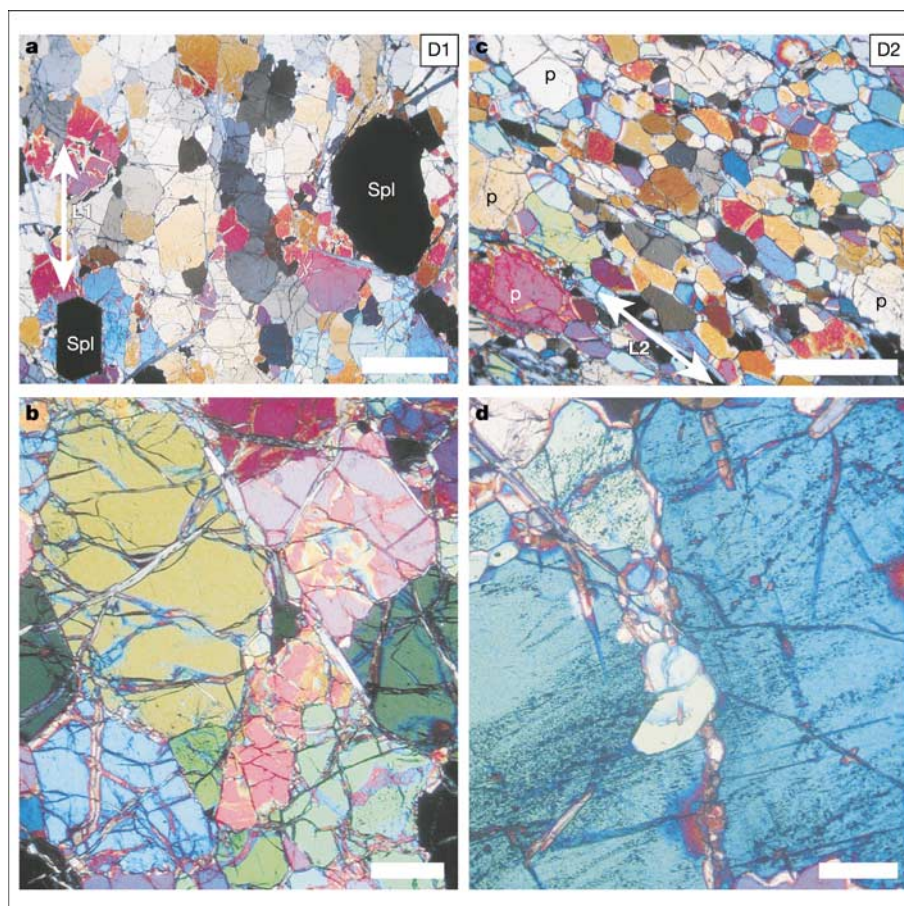


Figure 1 Photomicrographs of dunite in the Higashi-akaishi body (crossed polars). **a**, Coarse-grained D_1 dunite consisting of millimetre-sized olivine and Cr-spinel (Spl). A parallel alignment of their long axes defines a mineral lineation (L_1). Scale bar, 1.0 mm. **b**, Almost entirely inclusion-free D_1 olivine grains. Scale bar, 0.3 mm. **c**, Porphyroclastic

dunite consisting of coarse porphyroclasts (p) and D_2 neoblasts of olivine. The grain shape of the D_2 olivine defines an L_2 lineation. Scale bar, 1.0 mm. **d**, D_2 neoblasts formed along subgrain boundaries within porphyroclasts. The olivine porphyroclasts contain abundant micro-inclusions. Scale bar, 0.1 mm.

cracks can be explained as the result of volumetric changes when the water reacted with the host olivine crystal to form serpentine. The ubiquitous occurrence of these inclusions in the D_2 porphyroclasts and their absence in D_1 olivine imply, therefore, that D_2 took place in the presence of water, a necessary condition for B-type LPO to develop. The water was probably supplied by dehydration reactions, which characterize much of the metamorphic changes in subducted crustal rocks.

The depth at which the B-type LPO patterns were formed can be constrained using a combination of petrological and microstructural methods. The solubility of alumina in orthopyroxene coexisting with garnet is one of the best-established and widely applied methods of estimating the formation pressure of mantle rocks²⁴. Application of this geobarometer to the Higashi-akaishi body, combined with garnet-clinopyroxene and two-pyroxene geothermometers, has revealed a pressure-temperature evolution with a roughly isothermal burial at temperatures of 700–800 °C from a depth of 50 km to over 100 km (ref. 14). To relate this pressure-temperature history to the deformation stages and LPO patterns described above, we analysed the chemical compositions of orthopyroxene in a garnet peridotite in which both orthopyroxene and olivine have been dynamically recrystallized during the D_2 stage to form porphyroclasts and neoblasts. The Al content of

orthopyroxene is lower in domains affected by D_2 deformation, such as the neoblasts, rims of porphyroclasts and boudinaged porphyroclasts, than in the cores of the porphyroclasts (Fig. 4). The decrease in the Al content of orthopyroxene during D_2 represents an increase in pressure and corresponds to the burial part of the pressure-temperature path derived for the Higashi-akaishi body. We conclude, therefore, that the D_2 deformation and associated olivine LPO formed under increasing pressure at depths between 50 and 100 km.

Thus we have found regionally developed B-type olivine LPO patterns characterized by an a -axis maximum perpendicular to the stretching direction. Microtextural and petrological data show that these LPO patterns developed at mantle depths greater than 50 km. The location of the Higashi-akaishi peridotite body at the highest structural levels of a subduction-type metamorphic belt, the pressure-temperature path and peak metamorphic conditions all strongly suggest that this unit represents a sliver of wedge mantle originally located above a subducting slab. This implies that D_2 and

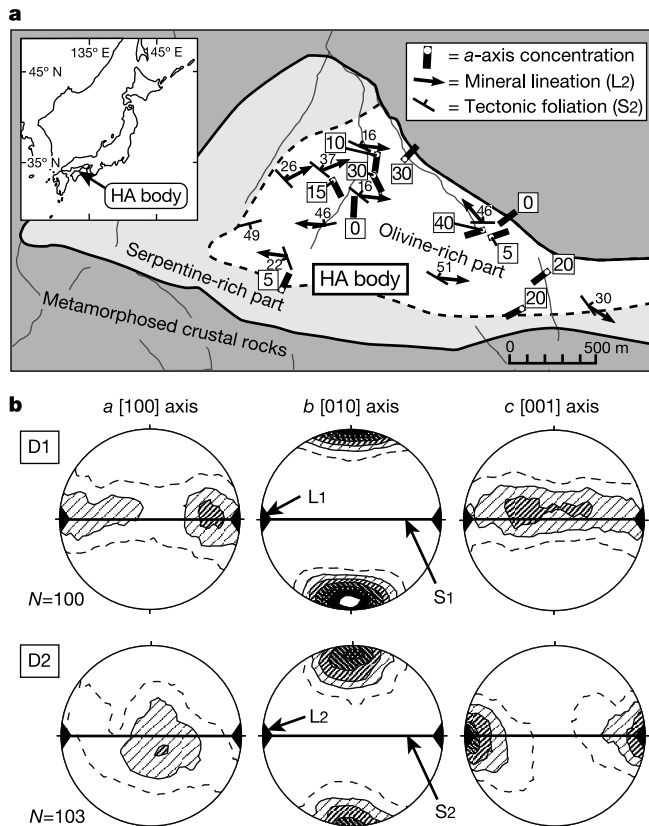


Figure 2 Structural analysis of olivine fabrics. **a**, Geological map of the Higashi-akaishi (HA) body showing representative D_2 structural data and the orientations of a -axis concentrations of the D_2 olivine LPO patterns. In each case the a -axis concentration is subperpendicular to the adjacent mineral lineation. Boxed numbers refer to the dips of the S_2 foliation and plunges of the a -axis concentrations. The LPO data include some measurements from ref. 16. **b**, Olivine LPO patterns of D_1 grains in coarse-grained dunite and of D_2 neoblasts in porphyroclastic dunite. The crystallographic axes are plotted with respect to the stretching lineation (L) and the tectonic foliation (S). N , number of measurements. Equal-area lower-hemisphere projection, with contours at 1 to 7 multiples of random concentration.

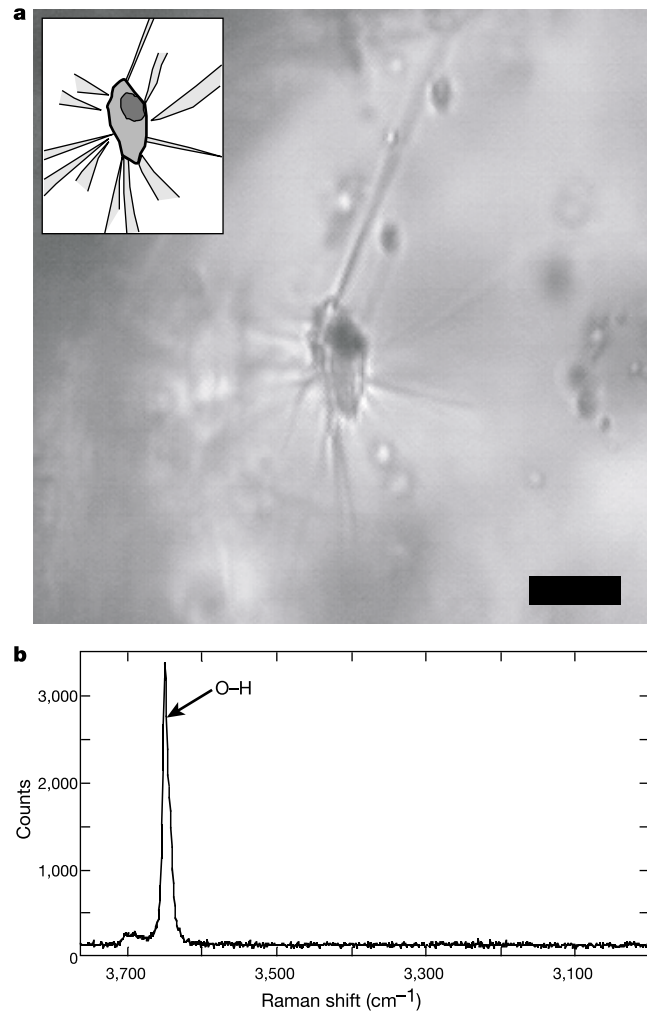


Figure 3 Micro-inclusions indicating the presence of water during the D_2 stage. **a**, Photomicrograph of micro-inclusions in olivine. An inclusion at the centre has a prismatic form, suggesting that it represents a negative crystal with the main fractures radiating from the apexes of the faceted surface. Raman spectroscopy shows that the light part of the inclusion and the surrounding fractures are filled by serpentine. Scale bar, 10 μ m. **b**, Raman spectrum for the light part of the inclusion in **a**. A sharp peak at the wavenumber of 3,650 cm^{-1} shows the presence of O-H bonds as water of crystallization²⁵.

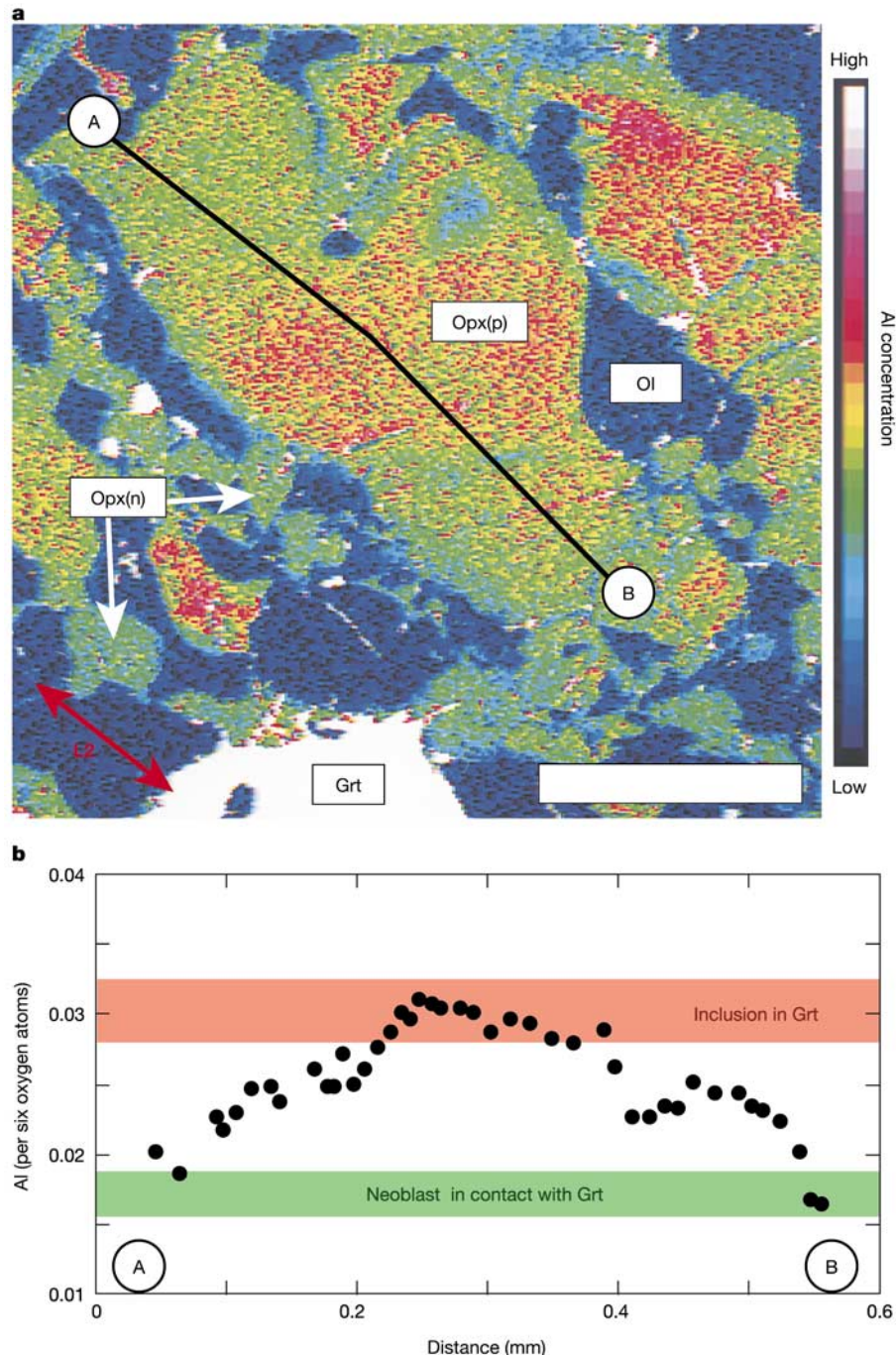


Figure 4 Mineral chemistry in garnet peridotite. **a**, X-ray elemental map using the Al K α line for minerals in garnet peridotite. Colours in the bar at the right side correspond to the X-ray intensity. The areas of relatively low Al (greenish) tend to occupy the extensional sides of the porphyroclastic orthopyroxene (Opx-p). Al concentration is also low in the

recrystallized neoblasts (Opx-n). In contrast, relatively high Al concentrations are found in the core. Ol, olivine; Grt, garnet. Scale bar, 0.2 mm. **b**, Line profile of Al content in the porphyroclastic orthopyroxene at the centre of **a**, determined using a microprobe.

the associated olivine LPO developed in subduction zone upper mantle.

Our results confirm the suggestion from experimental work¹⁴ that B-type olivine LPO patterns can develop in the mantle wedge above subducting slabs in the presence of water. In most subduction zones the mantle wedge is likely to be infiltrated with water and to experience similar physical conditions to those recorded here. We suggest, therefore, that mantle deformation above subduction zones should, in general, be associated with B-type olivine LPOs. This general association can explain the anomalous seismic anisotropy

in these regions and implies that mantle flow in these regions is parallel to plate motion, not perpendicular as has generally been thought^{10–13}. This new constraint on mantle flow is essential to build realistic physical models of subduction zones. □

Received 7 June; accepted 3 November 2003; doi:10.1038/nature02179.

1. Nicolas, A. & Christensen, N. I. Formation of anisotropy in upper mantle peridotites—a review. In *Composition, Structure and Dynamics of the Lithosphere-Asthenosphere System* (eds Fuchs, K. & Froidevaux, C.) 111–123 (1987).
2. Carter, N. L. & Ave'Lallemant, H. G. High temperature flow of dunite and peridotite. *Geol. Soc. Am. Bull.* **81**, 2181–2202 (1970).
3. Zhang, S. & Karato, S. Lattice preferred orientation of olivine aggregates deformed in simple shear.

- Nature* **375**, 774–777 (1995).
- Raitt, R. W., Shor, G. G., Francis, T. J. G. & Morris, G. B. Anisotropy of the Pacific upper mantle. *J. Geophys. Res.* **74**, 3095–3109 (1969).
 - Christensen, N. I. The magnitude, symmetry and origin of upper mantle anisotropy based on fabric analyses of ultramafic tectonites. *Geophys. J. R. Astron. Soc.* **76**, 89–111 (1984).
 - Hess, H. H. Seismic anisotropy of the uppermost mantle under oceans. *Nature* **203**, 629–631 (1964).
 - Francis, T. J. G. Generation of seismic anisotropy in the upper mantle along the mid-oceanic ridges. *Nature* **221**, 162–165 (1969).
 - Peselnick, L., Nicolas, A. & Stevenson, P. R. Velocity anisotropy in a mantle peridotite from the Ivrea zone: application to upper mantle anisotropy. *J. Geophys. Res.* **79**, 1175–1182 (1974).
 - Ando, M., Ishikawa, Y. & Yamazaki, F. Shear wave polarization anisotropy in the upper mantle beneath Honshu, Japan. *J. Geophys. Res.* **88**, 5850–5864 (1983).
 - Shih, X. R., Schneider, J. F. & Meyer, R. P. Polarities of P and S waves, and shear wave splitting observed from the Bucaramanga Nest, Colombia. *J. Geophys. Res.* **96**, 12069–12082 (1991).
 - Yang, X. & Fischer, K. M. Seismic anisotropy beneath the Shumagin Islands segment of the Aleutian-Alaska subduction zone. *J. Geophys. Res.* **100**, 18165–18177 (1995).
 - Fouth, M. J. & Fischer, K. M. Mantle anisotropy beneath northwest Pacific subduction zones. *J. Geophys. Res.* **101**, 15897–16002 (1996).
 - Smith, G. P. et al. A complex pattern of mantle flow in the Lau backarc. *Science* **292**, 713–716 (2001).
 - Jung, H. & Karato, S. Water-induced fabric transitions in olivine. *Science* **293**, 1460–1463 (2001).
 - Enami, M., Mizukami, T. & Yokoyama, K. Metamorphic evolution of garnet-bearing ultramafic rocks from the Gongen area, Sanbagawa belt, Japan. *J. Metamorph. Geol.* **22**, 1–15 (2004).
 - Yoshino, G. Structural-petrological studies of peridotite and associated rocks of the Higashi-akaishi-yama District, Shikoku, Japan. *J. Sci. Hiroshima Univ. C* **3**, 343–402 (1961).
 - Banno, S. & Sakai, C. Geology and metamorphic evolution of the Sanbagawa metamorphic belt, Japan. In *Evolution of Metamorphic Belts* (eds Daley, J. S., Cliff, R. A. & Yardley, B. W. D.) 519–532 (1989).
 - Takasu, A. & Dallmeyer, R. D. ⁴⁰Ar/³⁹Ar mineral age constraints for the tectonothermal evolution of the Sanbagawa metamorphic belt, central Shikoku, Japan: a Cretaceous accretionary prism. *Tectonophysics* **185**, 111–139 (1990).
 - Wallis, S. Exhuming the Sanbagawa metamorphic belt: the importance of tectonic discontinuities. *J. Metamorph. Geol.* **16**, 83–95 (1998).
 - Mori, T. & Banno, S. Petrology of peridotite and garnet clinopyroxenite of the Mt. Higashi Akaishi mass, central Shikoku, Japan—subsolidus relation of anhydrous phases. *Contrib. Mineral. Petrol.* **41**, 301–323 (1973).
 - Mercier, J.-C. C. & Nicolas, A. Textures and fabrics of upper-mantle peridotites as illustrated by xenoliths from basalts. *J. Petrol.* **16**, 454–487 (1975).
 - Boudier, F. & Nicolas, A. Nature of the Moho transition zone in the Oman ophiolite. *J. Petrol.* **36**, 777–796 (1995).
 - Hirai, H. & Arai, S. H₂O–CO₂ fluids supplied in alpine-type mantle peridotites: electron petrology of relic fluid inclusions in olivines. *Earth Planet. Sci. Lett.* **85**, 311–318 (1987).
 - Harley, S. L. & Green, D. H. Garnet-orthopyroxene barometry for granulites and peridotites. *Nature* **300**, 697–701 (1982).
 - Rossmann, G. R. Vibrational spectroscopy of hydrous components. *Rev. Mineral.* **18**, 193–206 (1988).

Acknowledgements S. Wallis aided in the field studies for this work and was closely involved in developing the rationale for the study. J. Yamamoto carried out the laser micro-Raman spectroscopic analyses. We thank H. Kagi for his support during these analyses and for discussion on this work. We also thank M. Enami and M. Obata for their comments. This work was supported in part by a grant from the Fukada Geological Institute.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.M. (mizukami@kueps.kyoto-u.ac.jp).

Friction falls towards zero in quartz rock as slip velocity approaches seismic rates

Giulio Di Toro¹, David L. Goldsby² & Terry E. Tullis²

¹Dipartimento di Geologia, Paleontologia e Geofisica, Università di Padova, Padova, 35137, Italy

²Department of Geological Sciences, Brown University, Providence, Rhode Island 02912, USA

An important unsolved problem in earthquake mechanics is to determine the resistance to slip on faults in the Earth's crust during earthquakes¹. Knowledge of coseismic slip resistance is critical for understanding the magnitude of shear-stress reduction and hence the near-fault acceleration that can occur during earthquakes, which affects the amount of damage that

earthquakes are capable of causing. In particular, a long-unresolved problem is the apparently low strength of major faults^{2–6}, which may be caused by low coseismic frictional resistance³. The frictional properties of rocks at slip velocities up to 3 mm s^{−1} and for slip displacements characteristic of large earthquakes have been recently simulated under laboratory conditions⁷. Here we report data on quartz rocks that indicate an extraordinary progressive decrease in frictional resistance with increasing slip velocity above 1 mm s^{−1}. This reduction extrapolates to zero friction at seismic slip rates of ~1 m s^{−1}, and appears to be due to the formation of a thin layer of silica gel on the fault surface: it may explain the low strength of major faults during earthquakes.

There are significant experimental difficulties in determining the resistance to slip on faults during earthquakes. Laboratory experiments need to combine the high slip velocities V (0.1–2 m s^{−1}), large slip displacements δ (0–10 m) and high normal stresses σ_n (>50 MPa) that might be necessary to activate dynamic fault weakening mechanisms operative during earthquakes. All existing laboratory friction data satisfy at most two of these three criteria^{7–13}. Values of the coefficient of friction μ for most rocks are relatively high over a wide range of normal stress; values of μ of ~0.6–0.85 are found when the slip velocity and displacement are ≤ 1 mm s^{−1} and <1 mm, respectively⁹. This high friction at ambient normal stresses in the Earth's continental crust is consistent with the magnitudes of shear stresses measured in the crust¹⁴. However, several mechanisms have been proposed that could lower shear resistance during fast coseismic slip, such as shear melting^{11,12,15,16}, pore fluid pressurization³, normal interface vibrations¹⁷, acoustic fluidization¹⁸ and elastohydrodynamic lubrication¹⁹. Given all of these potential dynamic weakening mechanisms, it seems quite plausible that resistance during earthquake slip might be lower than implied by the high values of friction⁹ measured at slow slip velocities. Nevertheless, these mechanisms and/or their applicability to earthquakes are still poorly understood; thus resistance to slip during earthquakes is still unknown.

To increase our understanding of coseismic slip resistance, we conducted rapid-slip experiments at ambient temperature and humidity in a servo-controlled compression–torsion apparatus. Sliding occurred during rotary shear of an annular surface oriented normal to the rotation axis. Experiments were performed on six samples of Arkansas novaculite (a quartz rock of meta-sedimentary origin²⁰ with a grain size of 1–5 μ m) and one sample of granite. The upper, ring-shaped part of the sample, with a thickness of 2.56 mm and inner and outer radii of 22.2 and 26.97 mm, respectively, was slid on a lower circular plate of ~35 mm radius and ~8 mm thickness. Sliding surfaces were pre-roughened with no. 150 SiC grit to create a r.m.s. surface roughness of ~30 μ m. Experiments were conducted at $\sigma_n = 5$ MPa. Temperatures were measured with a thermocouple embedded a distance of ~1 mm from the sliding surface in the upper sample ring. The rotary shear apparatus allows unidirectional sliding displacements of only ~40 mm; large cumulative displacements of up to 4,700 mm were achieved by repeatedly reversing the sliding direction (Fig. 1c). A constant slip velocity V in the range 1 μ m s^{−1} to 100 mm s^{−1} was maintained during each experiment (Fig. 1b–d).

The experiments involved three steps: an initial 'loading' step to verify the low-speed behaviour, a 'high-speed' step to determine the velocity dependence, and a final 'recovery' step to determine whether and how the strength returned to typical low-speed values (Fig. 1a). During the loading step, the sample was slid at 1 μ m s^{−1} for 1.2–2.0 mm of displacement. During this step, μ gradually increased from 0 to 0.7–0.8, then oscillated owing to 'stick-slip' sliding, as expected for the stiffness of the machine and the low-speed frictional behaviour of the rock⁸. In the high-speed step, V was abruptly increased and the sample was slid to large displacements, frequently up to ~4.5 m (Fig. 1a), during which friction decreased

dramatically. Then, in the recovery step, the sample was slid again at $1 \mu\text{m s}^{-1}$ for 1.2–2.0 mm of additional displacement, and friction gradually increased from high-speed values to low-speed values; this occurred over a time of $\sim 300 \text{ s}$ (right part of Fig. 1a).

The dependence of the frictional resistance on slip displacement during the high-speed step is characterized by a 'transient stage' and a 'steady-state stage' (Figs 1a and 2). In the transient stage, the value of μ decreases dramatically over ~ 0.5 –1 m of slip displacement towards a nearly constant value of 'steady-state friction' (μ_{ss}) that depends strongly on slip velocity. At our highest slip velocity of 100 mm s^{-1} , μ_{ss} is very low (0.2). During the steady-state stage, μ_{ss} decreases slightly with displacement at a rate of $\sim 0.0025 \pm 0.0005 \text{ m}^{-1}$.

The detailed nature of the dependence of friction on slip displacement during the transient stage is shown in Fig. 2. At 30 and 100 mm s^{-1} , the frictional sliding behaviour is quite reproducible, whereas at 3 mm s^{-1} the behaviour is more variable and appears to depend on the previous sliding history of the sample. The increased variability at 3 mm s^{-1} might be due to the fact that this velocity is near the transition from classic rate and state friction behaviour⁸, with its small velocity dependence of frictional strength, to behaviour dominated by another deformation mechanism at higher velocities. A temporary reduction and recovery in friction during the transient stage is often observed (Fig. 2b).

The magnitude of the dependence of μ_{ss} on slip velocity V is very marked; μ_{ss} decreases sharply between 3 and 100 mm s^{-1} (Fig. 3). Data in Fig. 3 are corrected for the small dependence of μ_{ss} on slip displacement ($\Delta\mu_{ss}/\Delta\delta = 0.0025 \text{ m}^{-1}$), and are presented for $\delta = 0$. For $V > 1 \text{ mm s}^{-1}$ the data can be described by $\mu_{ss} = 0.018(\pm 0.044) - [0.193(\pm 0.022) \log_{10} V]$ (where V is in m s^{-1}). The data extrapolate to a friction coefficient of zero at a sliding velocity of $\sim 1.2 \text{ m s}^{-1}$. Values of the friction coefficient shown in Fig. 3 are even lower than those observed in frictional sliding experiments on gabbro in which melting occurs¹¹ (filled black triangles in Fig. 3).

What mechanism is responsible for the extraordinary weakening observed in our experiments? Powdered rock (gouge) is produced continually and extruded from the sliding surface during our

experiments (the gouge layer was not confined laterally). After the experiments, the fault surface is covered by white 'flakes', $< 10 \mu\text{m}$ thick, of ultra-comminuted gouge. We presume there is a large amount of amorphous silica in these flakes, because examination via transmission electron microscopy of gouge flakes produced in friction experiments on quartz rocks with low slip displacements and sliding velocities shows that the gouge consists of quartz fragments and amorphous silica²¹. However, the amorphous silica alone cannot account for the extraordinary weakening, because at low sliding velocities the friction coefficient is 0.6–0.8.

To determine whether the temperature might have become high enough for melting to occur and perhaps explain the weakening, we conducted time-dependent finite element method calculations of the temperature and made direct thermocouple measurements. For the highest slip speed, 100 mm s^{-1} , the maximum temperature measured and calculated at 1 mm from the sliding interface is $\sim 100^\circ\text{C}$ and that calculated at the interface is $\sim 150^\circ\text{C}$. Frictional resistance may depend more critically on the locally higher temperatures at the small asperity contacts that actually support the normal and shear loads. These so-called 'flash temperatures' can be substantially higher than the average surface temperature²². An upper bound on the flash temperature of $\sim 400^\circ\text{C}$ is calculated by assuming that the contacts have a diameter of $10 \mu\text{m}$ and deform plastically at a contact stress of 10 GPa (refs 22, 23). As the melting point of cristobalite, the high-temperature quartz polymorph, is $1,713^\circ\text{C}$ at room pressure²⁴, neither the average nor flash temperatures are high enough to induce melting.

The weakening appears to be due to weak, fluid-like behaviour of a tribolayer of silica gel⁷, the production of which appears to require water from the humid atmosphere²⁵ and probably the presence of amorphous silica. If silica gel is responsible for the weakening, why is high slip velocity required for it to act as a weak fluid? We believe this is due to thixotropic behaviour²⁶ of the gel layer, in which shear-weakening is promoted by slip displacement, and time-dependent strengthening is inhibited by the short times available during high-velocity slip. If the gel is thixotropic, time-dependent recovery of shear strength is expected²⁶, consistent with the gradual recovery of shear strength with time during low-speed sliding immediately

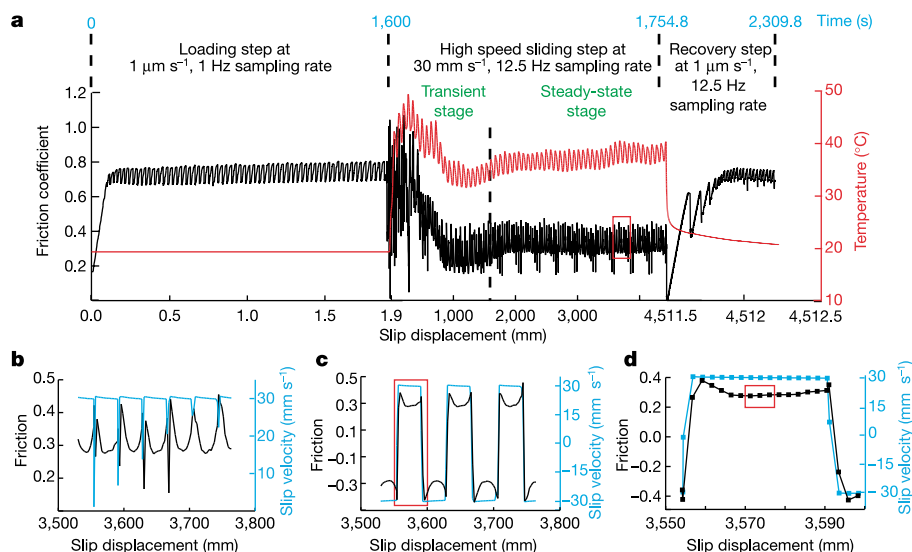


Figure 1 Typical experimental sequence and results. **a**, Each experiment consisted of three steps: a loading step, a high-speed sliding step, and a recovery step. The abscissa is expanded for the loading and recovery steps. The temperature is measured 1 mm from the sliding surface. Red box is expanded in **b**. **b**, Details of the high-speed step, with sliding velocity in blue. Positive and negative values of the friction coefficient, recorded in

opposite sliding directions, have been rectified so that friction is plotted as positive values in **a** and **b**. The many vertical spikes in the rectified friction coefficient data (**a**, **b**) occur at reversals of the sliding direction, as shown in the unrectified data of **c**. Red box in **c** is expanded in **d**. **d**, Details of one reversal. The friction (as plotted in Figs 2 and 3) was measured between the reversals of sliding direction, as shown in the red box in **d**.

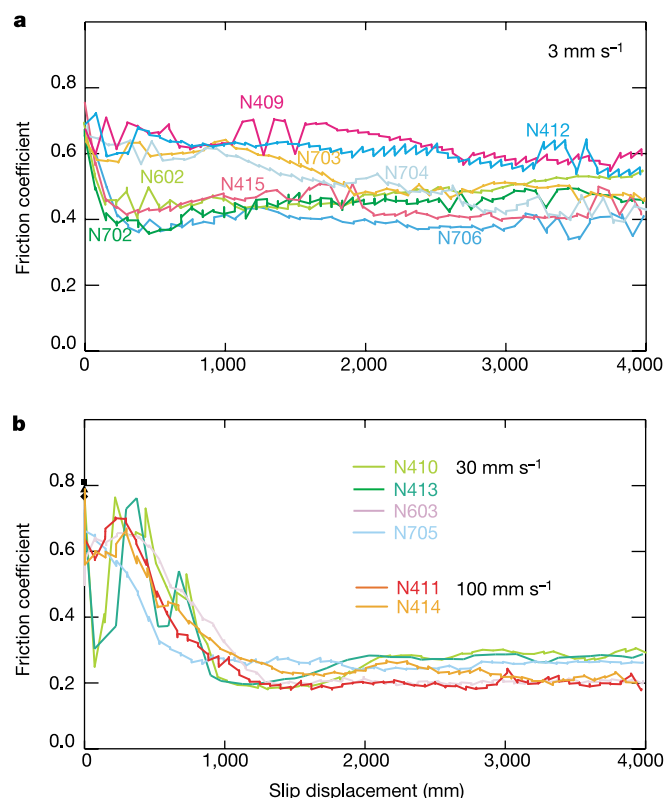


Figure 2 Dependence of friction on slip displacement for samples slid at 3, 30 and 100 mm s⁻¹. The friction data have been filtered to remove the spikes associated with the reversals in direction, as shown in Fig. 1d. The letter N and first digit of the data label corresponds to the particular novaculite sample tested; the last two digits are a sequential number showing the order in which the data for successive experiments on that sample were collected. **a**, 3 mm s⁻¹ data. At this speed, there is considerable variability in the frictional behaviour; this variability depends in part on the previous sliding history of a sample. After sliding at 30 or 100 mm s⁻¹, the subsequent behaviour at 3 mm s⁻¹ is consistently seen to fall off within δ of <500 mm. (N412 is not an exception; it followed N410, but the sample surface was exposed and cleaned of wear debris before N412). **b**, 30 and 100 mm s⁻¹ data. At these velocities, friction typically undergoes a temporary reduction of variable magnitude and returns to nearly the original value within the first 200–400 mm of slip, followed by a large reduction towards the steady-state value. The square, triangle and diamond at $\delta = 0$ show the initial friction (at $V = 1 \mu\text{m s}^{-1}$) for samples N400, N600 and N700, respectively.

following high-speed sliding seen in the recovery step (right part of Fig. 1a). In this study, the recovery of frictional strength has a time constant of ~ 100 s; the fact that the strength recovery is not instantaneous rules out the other weakening mechanism that could be caused by a weak fluid, namely elastohydrodynamic lubrication¹⁹.

The temporary reduction and recovery of strength that we typically observe during the transient stage (Fig. 2b) is similar to that seen in high-speed experiments in which shear melting occurs, and for which flash melting, followed by freezing and asperity welding, is offered as the explanation¹¹. Perhaps local production of gel at contacts could similarly be responsible for our temporary weakening, although an explanation for the restrengthening and subsequent weakening to the steady state is not obvious.

The importance of this weakening mechanism for earthquakes is unclear at present. For example, the amount of silica required in a rock to produce a weak layer of silica gel is not yet known. As shown in Fig. 3, granite (with less silica) shows nearly constant frictional strength as V increases from 3 to 30 mm s⁻¹ at $\sigma_n = 5$ MPa (as does gabbro¹¹), but at $\sigma_n = 112$ MPa the friction coefficient of granite at

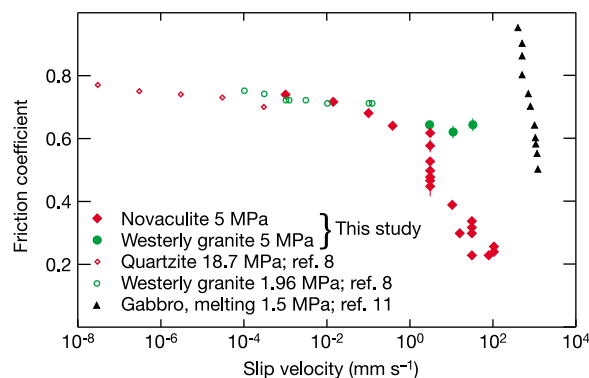


Figure 3 Dependence of steady-state friction on slip velocity. Data for novaculite for $V > 1 \text{ mm s}^{-1}$ were collected at δ between 3 and 4 m and corrected to $\delta = 0$ for the slight trend of μ_{ss} with δ ($0.0025 \pm 0.0005 \text{ m}^{-1}$). Data from this study (filled green and red symbols) for novaculite and granite are compared with data from the literature. At V around 1–3 mm s⁻¹, novaculite shows an abrupt decrease in shear resistance, and attains values of μ even lower than those attained during frictional melting of gabbro. For our data, the standard deviation for each individual data point (namely, the standard deviation for the corresponding curve in Fig. 2 for δ between 3 and 4 m) is either smaller than the plotted symbol or is shown by a short vertical bar.

3.2 mm s⁻¹ is less than 0.4, so weakening due to silica gel formation may occur in granite at higher σ_n (ref. 7). We note that many seismic faults are mineralized with quartz even when the country rock is not quartz-bearing²⁷, suggesting that quartz behaviour could be important for earthquakes in many rock types. We also note that pseudotachylites are rarely found in rocks composed primarily of quartz, whereas they are more abundant in rocks with a higher content of mafic minerals²⁸, as expected if weakening owing to gel formation prevents the large degree of frictional heating needed to cause melting.

This dynamic weakening mechanism—or others that also require a significant amount of slip before weakening occurs and a comparatively large amount of time to restrengthen after rapid slip—could yield interesting behaviour not embodied in current models of earthquake source dynamics. For example, strong velocity weakening behaviour is usually found to be a necessary condition for earthquake rupture to occur as self-healing slip pulses^{29,30}, but if strength recovery occurs over times of 100 s or more, such pulses would seem unlikely, as healing would not occur until all coseismic slip has ceased. With or without self-healing slip pulses, the dramatic reduction in shear stress with slip velocity that we observe could allow slip to occur during earthquakes with virtually no resistance. □

Received 19 July; accepted 14 November 2003; doi:10.1038/nature02249.

1. Kanamori, H. Mechanics of earthquakes. *Annu. Rev. Earth Planet. Sci.* **22**, 207–237 (1994).
2. Brune, J. N., Henyey, T. L. & Roy, R. F. Heat flow, stress, and rate of slip along the San Andreas fault, California. *J. Geophys. Res.* **74**, 3821–3827 (1969).
3. Lachenbruch, A. H. Frictional heating, fluid pressure, and the resistance to fault motion. *J. Geophys. Res.* **85**, 6249–6272 (1980).
4. Mount, V. S. & Suppe, J. State of stress near the San Andreas fault: Implications for wrench tectonics. *Geology* **15**, 1143–1146 (1987).
5. Zoback, M. D. et al. New evidence on the state of stress of the San Andreas fault system. *Science* **238**, 1105–1111 (1987).
6. Rice, J. R. in *Fault Mechanics and Transport Properties of Rocks* (eds Evans, B. & Wong, T.) 475–503 (Academic, San Diego, 1992).
7. Goldsby, D. & Tullis, T. E. Low frictional strength of quartz rocks at subseismic slip rates. *Geophys. Res. Lett.* **29**, doi:10.1029/2002GL015240 (2002).
8. Dieterich, J. H. Time-dependent friction and the mechanics of stick slip. *Pure Appl. Geophys.* **116**, 790–806 (1978).
9. Byerlee, J. D. Friction of rocks. *Pure Appl. Geophys.* **116**, 615–626 (1978).
10. Tullis, T. E. & Weeks, J. D. Constitutive behavior and stability of frictional sliding of granite. *Pure Appl. Geophys.* **124**, 383–414 (1986).
11. Tsutsumi, A. & Shimamoto, T. High-velocity frictional properties of gabbro. *Geophys. Res. Lett.* **24**, 699–702 (1997).
12. Spray, J. G. Artificial generation of pseudotachylite using friction welding apparatus: simulation of

- melting on a fault plane. *J. Struct. Geol.* **9**, 49–60 (1987).
13. Mair, K. & Marone, C. Shear heating in granular layers. *Pure Appl. Geophys.* **157**, 1847–1866 (2000).
 14. Hickman, S. H. Stress in the lithosphere and the strength of active faults. *Rev. Geophys.* **29**, 759–775 (1991).
 15. McKenzie, D. P. & Brune, J. N. Melting of fault planes during large earthquakes. *Geophys. J. R. Astron. Soc.* **29**, 65–78 (1972).
 16. Spray, J. G. Viscosity determinations of some frictionally generated silicate melts — implications for fault zone rheology at high-strain rates. *J. Geophys. Res.* **98**, 8053–8068 (1993).
 17. Brune, J. N., Brown, S. & Johnson, P. A. Rupture mechanism and interface separation in foam rubber models of earthquakes: a possible solution to the heat flow paradox and the paradox of large overthrusts. *Tectonophysics* **218**, 59–67 (1993).
 18. Melosh, H. J. Dynamic weakening of faults by acoustic fluidization. *Nature* **397**, 601–606 (1996).
 19. Brodsky, E. E. & Kanamori, H. Elastohydrodynamic lubrication of faults. *J. Geophys. Res.* **106**, 16357–16374 (2001).
 20. Knauth, L. P. in *Silica: Physical Behavior, Geochemistry, and Materials Applications* (eds Heaney, P. J., Prewitt, C. T. & Gibbs, G. V.) 233–258 (Mineralogical Society of America, Washington DC, 1994).
 21. Yund, R. A., Blanpied, M. L., Tullis, T. E. & Weeks, J. D. Amorphous material in high strain experimental fault gouges. *J. Geophys. Res.* **95**, 15589–15602 (1990).
 22. Archard, J. F. The temperature of rubbing surfaces. *Wear* **2**, 438–455 (1958).
 23. Dieterich, J. H. & Kilgore, B. D. Imaging surface contacts: power law contact distributions and contact stresses in quartz, calcite, glass, and acrylic plastic. *Tectonophysics* **256**, 219–239 (1996).
 24. Deer, W. A., Howie, R. A. & Zussman, J. *An Introduction to the Rock Forming Minerals* 2nd edn (Longman Scientific and Technical, Harlow, UK, 1992).
 25. Titone, B., Sayre, K., Di Toro, G., Goldsby, D. L. & Tullis, T. E. The role of water in the extraordinary frictional weakening of quartz rocks during rapid sustained slip. *Eos* **82**, T31B–0841 (2001).
 26. Ferguson, J. & Kemblowski, Z. *Applied Fluid Rheology* (Elsevier, New York, 1991).
 27. Power, W. L. & Tullis, T. E. The relationship between slickenside surfaces in fine-grained quartz and the seismic cycle. *J. Struct. Geol.* **11**, 879–893 (1989).
 28. Killick, A. M. & Roering, C. An estimate of the physical conditions of pseudotachylite formation in the West Rand Goldfield, Witwatersrand Basin, RSA. *Tectonophysics* **284**, 247–259 (1998).
 29. Perrin, G. O., Rice, J. R. & Zheng, G. Self-healing slip pulse on a frictional surface. *J. Mech. Phys. Solids* **43**, 1461–1495 (1994).
 30. Beeler, N. M. & Tullis, T. E. Self-healing slip pulses in dynamic rupture models due to velocity dependent strength. *Bull. Seismol. Soc. Am.* **86**, 1130–1148 (1996).

Acknowledgements We thank C. Bull for assistance with the experimental apparatus, D. McGuirl for fabrication of experimental parts, M. Parmentier for assistance with the FEM program, K. Mair for suggestions about Fig. 1, J. Rice and N. Beeler for discussions throughout the course of this study, and J. Tullis and C. Marone for comments on the manuscript. This research was supported by the US Geological Survey, the US National Science Foundation, and MURST (Italy).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to T.E.T. (Terry_Tullis@brown.edu).

A new orang-utan relative from the Late Miocene of Thailand

Yaowalak Chaimanee¹, Varavudh Suteethorn¹, Pratueng Jintasakul², Chavalit Vidthayanon³, Bernard Marandat⁴ & Jean-Jacques Jaeger⁴

¹Paleontology Section, Geological Survey Division, Department of Mineral Resources, Bangkok-10400, Thailand

²Rajabhat Institute Nakorn Ratchasima, Nakorn Ratchasima-30000, Thailand

³Museum and Aquarium Division, Department of Fisheries, Bangkok-10900, Thailand

⁴Paléontologie, I.S.E.M., cc 064, CNRS-Université Montpellier II, Place Eugene Bataillon, 34095 Montpellier, France

The fossil record of the living great apes is poor. New fossils from undocumented areas, particularly the equatorial forested habitats of extant hominoids, are therefore crucial for understanding their origins and evolution¹. Two main competing hypotheses have been proposed for orang-utan origins: dental similarities^{2,3} support an origin from *Lufengpithecus*, a South Chinese⁴ and Thai Middle Miocene hominoid²; facial and palatal similarities⁵ support an origin from *Sivapithecus*, a Miocene hominoid from the Siwaliks of Indo-Pakistan^{4,6}. However, materials other than teeth and faces do not support these hypotheses^{7,8}. Here we describe the lower jaw of a new hominoid from the Late Miocene

of Thailand, *Khoratpithecus piriyai* gen. et sp. nov., which shares unique derived characters with orang-utans and supports a hypothesis of closer relationships with orang-utans than other known Miocene hominoids. It can therefore be considered as the closest known relative of orang-utans. Ancestors of this great ape were therefore evolving in Thailand under tropical conditions similar to those of today, in contrast with Southern China and Pakistan, where temperate⁹ or more seasonal¹⁰ climates appeared during the Late Miocene.

Khoratpithecus piriyai gen. et sp. nov. corresponds to undistorted mandibular corpora, lacking part of the left side inferiorly and also right and left rami, preserving roots of left canine–right I₂ as well as the rest of the dentition. It originates from a sand pit in Chalerm Prakieat District, Nakorn Ratchasima Province (Khorat) of Northeastern Thailand (15° 01' 35" N, 102° 16' 50" E). Sediments correspond to large meandering channels with many tree trunks identified as cf. *Corypha* (palm), cf. *Terminalia*, cf. *Parashorea* and Dipterocarps (C. Vozenin-Serra, personal communication). Large mammals including proboscidiens (*Deinotherium indicum*, *Gomphotherium* sp., *Stegolophodon* sp. and primitive *Stegodon*), anthracothere, pig (*Hippopotamodon sivalensis*), rhinos (*Chilotherium palaeosinense*, *Brachypotherium perimense* and *Alicornops complanatum*), bovids, giraffids (*Sivatherium* sp.) and rare hipparions are present. The faunal assemblage corresponds to Upper Nagri and Lower Dhok Pathan Formations of Siwaliks and indicates an early Late Miocene age, between 9 and 7 Myr (million years)¹¹.

Systematics. Order Primates Linnaeus 1758; suborder Anthropoidea Mivart 1864; superfamily Hominoidea Gray 1825; family Hominidae Gray 1825; subfamily Ponginae Elliot 1913; *Khoratpithecus* gen. nov.

Type species. *Khoratpithecus piriyai* sp. nov.

Referred species. cf. *Lufengpithecus chiangmuanensis* Chaimanee et al. 2003, Middle Miocene of Thailand.

Etymology. *Khoratpithecus* means ape from Khorat.

Diagnosis. Large hominoid of the size of *Lufengpithecus*, with a high and very thick lower jaw corpus bearing a strong lateral eminence at M₃ level, broad canine–incisor area, U-shaped dental arcade. Symphysis long and proclined with weak upper transverse torus, shallow posteriorly oriented genial fossa and strong inferior transverse torus. Absence of anterior digastric muscle impressions. Canine buccolingually compressed with flat lingual surface, distal groove and shallow lingual cingulum. Proclined and large lateral incisor roots. Sectorial P₃ without metaconid cusp. P₄ with high talonid bearing strong cusps. M₃ large with discontinuous buccal cingulid, buccally located hypoconulid and large distal fovea. Its original character association distinguishes it from other Miocene hominoids. Differs also from *Lufengpithecus* by its larger anterior dentition, canine structure and larger M₃. From *Gigantopithecus* by its less distally extended symphysis, larger canine, sectorial P₃, less elongated molars and less distally divergent tooth rows. From *Ouranopithecus* by its sectorial P₃ and thinner enamel. From *Pongo* by its thicker symphysis with less distally elongated lower transverse torus, smaller incisor alveola, shorter P₄, less peripheralized molar cusps, M₃/M₂ proportions, less wrinkled enamel and stronger lateral eminence.

Khoratpithecus piriyai sp. nov.

Etymology. In honour of Piriya Vachajitpan, who played a critical part in recovering the fossil.

Holotype. Incomplete mandibular corpora, RIN 765 (Rajabhat Institute, Nakorn Ratchasima) (Fig. 1, Table 1).

Locality and age. Sand pit in Nakorn Ratchasima Province (Khorat), Northeastern Thailand, Late Miocene age.

Diagnosis. Differs from *Khoratpithecus chiangmuanensis* (Chaimanee et al. 2003) by its buccolingually compressed canine with flat lingual surface, larger lower incisors roots and enlarged M₃.

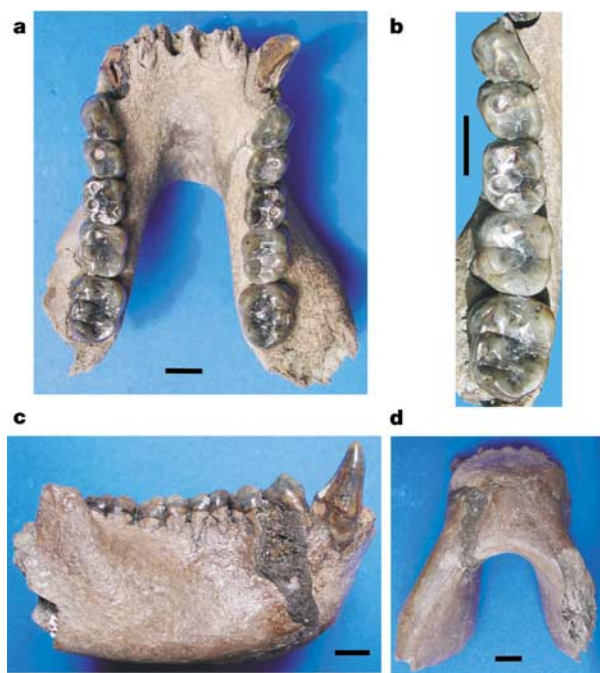


Figure 1 Mandible of *Khoratpithecus piriyai* gen. et sp. nov. holotype (RIN 765). **a, b**, Occlusal view. **c**, Lateral view. **d**, Inferior view. Scale bar, 1 cm.

Description. In occlusal view, the mandible displays a broad U-shaped dental arcade, widest at the level of canines and M_3 , with buccally concave tooth rows. Procumbent incisor roots are arranged on a slightly convex arcade and occupy a wide area (27 mm). Left canine section has a triangular outline with rounded corners. The right canine crown has a length/breadth ratio of 1.49, higher than in most other Miocene and extant apes excluding *Dryopithecus* and *Ouranopithecus*¹². Its relative height index is 1.48, less than that of *Kenyapithecus wickeri* (1.62), *Dryopithecus* (1.54–1.61) and *Lufengpithecus* (1.56–1.85), falling into the range of *Proconsul* and great apes (1.42–1.58) (ref. 13). The short and weak anterior ridge curves lingually into the cingulid and the posterior ridge ends into a talonid cusp. A wide groove develops buccally to the distal ridge. Lingual wall is flattened, as in *Griphopithecus*¹⁴, and bears two furrows bounding a medial ridge. A shallow lingual cingulid is present above the enamel–dentine junction. Canines of *K. chiangmuanensis* and *L. lufengensis* display rounded/oval sections at the dentine–enamel junction and *L. lufengensis* has a stronger lingual cingulid. The canine of *Sivapithecus* is less outwardly oriented, with more rounded crown and stronger lingual cingulid¹⁵. *K. piriyai* canine resembles those of some Pleistocene *Pongo*¹⁶, sharing flat lingual walls, wide distal grooves and shallow lingual cingulids. Premolars and M_1 – M_2 are similar to those of *K. chiangmuanensis* but molars differ from those of *L. lufengensis* by M_2/M_1 crown surface ratio (135% on *K. piriyai* and 172% on *L. lufengensis*). Regression of M_1 dimensions on body mass in extant primates^{17–19} indicates a body weight of approximately 70–80 kg. M_2 is broad, having a mesiodistal length/mesial width index of 108.5 ($N = 2$) that falls into the range of *Sivapithecus*²⁰. M_3 is very large, differing from *Lufengpithecus* by its M_3/M_1 crown surface ratio (196.5% for *K. piriyai*, 142% for male *L. lufengensis*, 128% for *L. keiyuanensis*) and bears a discontinuous buccal cingulid. Its central fovea is deep and displays enamel wrinkles. Talonid and distal fovea are large. It resembles *Ouranopithecus* by its buccal cingulid and on-line positioned buccal cusps. M_3 of *Gigantopithecus giganteus* has shallow central fovea and more elongated outline. Inferiorly, the mandible displays a long and wide symphyseal region. Medially, facets for the geniohyoid muscles are well expressed. Laterally to them, the basal

Table 1 Measurements of *Khoratpithecus piriyai* gen. et sp. nov. (RIN 765)

	Mesiodistal (mm)	Buccolingual	
		Trigonid (mm)	Talonid (mm)
C_R	13.71	9.33	
C_L	(13.78)	(10.11)	
P_{3L}	13.79	8.51	
P_{3R}	14.57	8.53	
P_{4L}	9.61	10.63	10.53
P_{4R}	9.24	10.65	10.23
M_{1L}	12.88	10.53	11.30
M_{1R}	12.38	10.69	11.34
M_{2L}	13.76	12.86	12.88
M_{2R}	13.77	12.68	12.72
M_{3L}	17.73	14.68	13.16
M_{3R}	17.63	14.71	13.12
C– M_3 length			78.90
P_3 – M_3 length			64.30
M_1 – M_3 length			42.97
C crown height (buccal)	20.40		
Corpus at P_4 height	38.88		
Corpus at P_4 thickness	21.23		
Corpus at M_3 height	36.33		
Corpus at M_3 thickness	31.43		
Symphysis height	47.14		
Symphysis thickness	21.00		
Maximal symphyseal length (infradentale–gnathion)	54.40		
Minimum symphyseal length (horizontal projection)	45.65		
External bi- P_3 breadth	54.42		
External bi- M_1 breadth	52.28		
External bi- M_3 breadth	58.95		
Breadth across incisor region	27.03		
Inter canine breadth	31.35		
Depth of mental foramen below alveolar margin	24.20		

R, right; L, left. Measurements in parentheses are estimated values.

symphyseal surface displays no impression of muscle insertion that could correspond to the anterior digastric muscles^{7,21}. Other Miocene hominoids display impressions of these muscles and their absence is considered as an exclusive specialization of extant *Pongo*, in relation to the development of laryngeal air sac^{7,21}. Laterally, the corpus is deep, with thickness/height index of 55% at P_4 and 87% at M_3 . Depth is nearly constant from symphysis to M_3 . Corpus shows a deep post- CP_3 depression, and an extreme thickness with strong lateral eminence at the M_3 level. The foramen mentale is located below mesio-buccal P_3 root, in a low position. The symphysis is strongly proclined (Fig. 2). In mid-sagittal cross-section, its long axis forms an angle of 42° to the alveolar margin of P_3 – M_3 , which is less than in other Miocene hominoids except *Equatorius* and *Kenyapithecus wickeri*²⁰. It falls in the range of extant large-bodied apes, but at the uppermost limit of adult *Pongo*²⁰. Its morphology and forward inclination are similar to those of *Griphopithecus*¹⁴, whose symphysis extends further distally with a deeper genial fossa. Internally, there is a long planum alveolare sloping at about 35° on the alveolar plane to a weak superior transverse torus. A shallow genial fossa, oriented distally, lies below the superior transverse torus, above a thick, rounded, inferior transverse torus that projects until the level of M_1 trigonid. The buccal symphyseal surface is wide and flat. The symphysis of *Ankarapithecus*¹², *Sivapithecus*²¹, *Gigantopithecus giganteus*, *Ouranopithecus*²² and *Lufengpithecus*²¹ are more vertically oriented (Fig. 2). *G. giganteus* and *Lufengpithecus* have no strongly developed inferior transverse tori, and *Sivapithecus* displays tori of relatively equal prominence^{7,21}. *Equatorius* displays a thicker symphysis section, with a stronger superior torus²⁰. *Pongo* shows stronger superior transverse torus, thinner section and more distally extended lower torus^{21,23,24}. However, some variants of extant orang-utan symphysis²¹ are similar to that of *Khoratpithecus*.

Several Miocene Asian hominoids have been proposed as potential ancestors of orang-utans. On the basis of its facial and palatal similarities⁵, *Sivapithecus* was considered as the best candidate. However, its lower jaw anatomy has been considered markedly dissimilar in all essential components to the *Pongo* mandible^{6,21}.

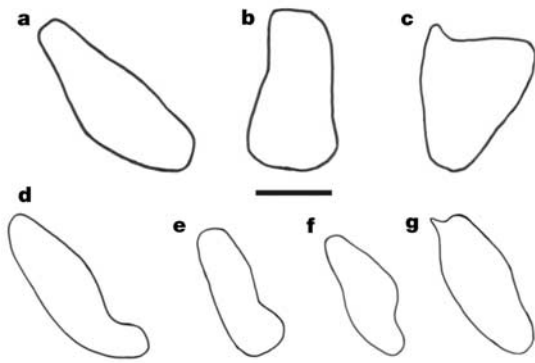


Figure 2 Mandibular section of *Khoratpithecus piriyai*. **a**, Symphyseal section. **b**, Coronal section at the middle of M₁. **c**, Coronal section at the middle of M₃. Scale bar, 2 cm. **d–g**, Symphyseal sections of *Pongo* (**d**), *Lufengpithecus* (**e**), *Sivapithecus* (**f**) and *Gigantopithecus* (**g**), not to scale²¹.

Lufengpithecus, from the Late Miocene of South China, displays greater tooth resemblances to *Pongo* than *Sivapithecus*^{4,25}, but it has been excluded from *Pongo* ancestry because of its different face and periorbital region structures. It has been allocated, like *Ankarapithecus*²⁶, to the sister taxon of the Ponginae + Homininae²⁶ or to a primitive sister taxon of the Ponginae^{4,27}. *K. chiangmuanensis*, from the Middle Miocene of Thailand², shares similar premolar and molar characters and large incisor roots with *K. piriyai*. Therefore we refer the Middle Miocene Thai species to the same new genus. Both species nevertheless differ significantly enough to be distinguished at specific level. *Ankarapithecus*, *Sivapithecus*, *Lufengpithecus* and *Gigantopithecus* are united as members of the *Pongo* clade^{26,27}. *Khoratpithecus* displays all characters that distinguish the members of that clade and an original character assemblage that occurs in none of the known Miocene hominoids, justifying its allocation to a new genus. Resemblance of *Khoratpithecus* premolars and molars to those of *Lufengpithecus* are interpreted as shared primitive characters of the orang-utan clade. *Khoratpithecus* jaw shares several derived characters with Pleistocene and extant orang-utans; among them, the absence of anterior digastric muscles represents an exclusive synapomorphy^{7,21} demonstrating that it corresponds to orang-utan's closest related fossil hominoid. Both *Khoratpithecus* species are found in association with tropical floras, indicating that extant apes were differentiating in areas where tropical conditions have prevailed from Middle to Late Miocene. In Southern China, Late Miocene *Lufengpithecus*⁴ are associated with a more temperate flora⁹. In Pakistan, Late Miocene *Sivapithecus* remains belong to mammal communities, indicating more open environments and increasing seasonality^{10,11}. Further discoveries in yet undocumented tropical areas seem to be crucial to an understanding of the origins and evolution of both orang-utans and African great apes. □

Received 10 October; accepted 11 November 2003; doi:10.1038/nature02245.

- Pilbeam, D. R. Genetic and morphological records of the Hominoidea and hominoid origins: A synthesis. *Mol. Phylog. Evol.* **5**, 155–168 (1996).
- Chaimanee, Y. et al. A new middle Miocene hominoid from Thailand and orangutan origins. *Nature* **422**, 61–65 (2003).
- Schwartz, J. H. in *Function, Phylogeny, and Fossils: Miocene Hominoid Evolution and Adaptations* (eds Begun, D. R., Ward, C. V. & Rose, M. D.) 363–388 (Plenum, New York, 1997).
- Kelley, J. in *The Primate Fossil Record* (ed. Hartwig, W. C.) 369–384 (Cambridge Univ. Press, 2002).
- Pilbeam, D. New hominoid skull material from the Miocene of Pakistan. *Nature* **295**, 232–234 (1982).
- Ward, S. in *Function, Phylogeny, and Fossils: Miocene Hominoid Evolution and Adaptations* (eds Begun, D. R., Ward, C. V. & Rose, M. D.) 269–290 (Plenum, New York, 1997).
- Brown, B. *The mandibles of Sivapithecus*. Thesis, Kent State Univ. Graduate Coll., Kent, Ohio (1989).
- Pilbeam, D. R., Rose, M. D., Barry, J. C. & Shah, S. M. I. New *Sivapithecus* humeri from Pakistan and the relationship of *Sivapithecus* and *Pongo*. *Nature* **348**, 237–239 (1990).
- Badgley, C., Guoqing, Q., Wanyong, C. & Defen, H. Paleocology of a Miocene, upland fauna: Lufeng, China. *Natl Geogr. Res.* **4**, 178–195 (1988).
- Morgan, M. E., Kingston, J. D. & Marino, B. D. Carbon isotopic evidence for the emergence of C₄

plants in the Neogene from Pakistan and Kenya. *Nature* **367**, 162–165 (1994).

- Barry, J. C. et al. Faunal and environmental change in the Late Miocene Siwaliks of Northern Pakistan. *Paleobiol. Mem.* **3**(28), 1–71 (2002).
- Begun, D. R. & Güleç, E. Restoration of the type and palate of *Ankarapithecus metei*: Taxonomic and phylogenetic implications. *Am. J. Phys. Anthropol.* **105**, 279–314 (1998).
- Ward, S., Brown, B., Hill, A., Kelley, J. & Downs, W. *Equatorius*: A new hominoid genus from the Middle Miocene of Kenya. *Science* **285**, 1382–1386 (1999).
- Alpagut, B., Andrews, P. & Martin, L. New hominoid specimens from the Middle Miocene site at Pasalar, Turkey. *J. Hum. Evol.* **19**, 397–422 (1990).
- Pilbeam, D., Rose, M. D., Badgley, C. & Lipschutz, B. Miocene hominoids from Pakistan. *Postilla* **181**, 1–94 (1980).
- Hooijer, D. A. Prehistoric teeth of man and of the orang-utan from Central Sumatra, with notes on the fossil orang-utan from Java and Southern China. *Zool. Med.* **29**, 175–301 (1948).
- Conroy, G. C. Problems of body-weight estimation in fossil primates. *Int. J. Primatol.* **8**, 115–137 (1987).
- Gingerich, P. D., Smith, H. S. & Rosenberg, K. Allometric scaling in the dentition of primates and prediction of body weight from tooth size in fossils. *Am. J. Phys. Anthropol.* **58**, 81–100 (1982).
- Legendre, S. Les communautés de mammifères du Paléogène (Éocène supérieur et Oligocène) d'Europe occidentale: structures, milieux et évolution. *Münch. Geowiss. Abh.* **16**, 1–110 (1989).
- McCrossin, M. L. & Benefit, B. R. Recently recovered *Kenyapithecus* mandible and its implications for great ape and human origins. *Proc. Natl Acad. Sci. USA* **90**, 1962–1966 (1993).
- Brown, B. in *Function, Phylogeny, and Fossils: Miocene Hominoid Evolution and Adaptations* (eds Begun, D. R., Ward, C. V. & Rose, M. D.) 153–171 (Plenum, New York, 1997).
- de Bonis, L. & Melentis, J. Les primates hominoides du Vallésien de Macédoine (Grèce). Étude de la machoire inférieure. *Géobios* **10**, 849–885 (1977).
- Daegling, D. J. Shape variation in the mandibular symphysis of apes: an application of a median axis method. *Am. J. Phys. Anthropol.* **91**, 505–516 (1993).
- Daegling, D. J. & Jungers, W. L. Elliptical Fourier analysis of symphyseal shape in great ape mandibles. *J. Hum. Evol.* **39**, 107–122 (2000).
- Harrison, T., Ji, X. & Su, D. On the systematic status of the late Neogene hominoids from Yunnan Province. *China. J. Hum. Evol.* **43**, 207–227 (2002).
- Alpagut, B. et al. A new specimen of *Ankarapithecus metei* from the Sinap Formation of central Anatolia. *Nature* **382**, 349–351 (1996).
- Begun, D. R. in *The Primate Fossil Record* (ed. Hartwig, W. C.) 339–368 (Cambridge Univ. Press, 2002).

Acknowledgements We thank L. de Bonis and D. Pilbeam for comments, help, discussions and for providing documents and comparative materials; P.-O. Antoine for the identification of large mammals; C. Vozenin-Serra for the identification of fossil wood; and M. Ponce de Léon and C. Zollikofer for the CT-Scan sections. This work is supported by the Fyssen and Leakey foundations, the Department of Mineral Resources (Bangkok), the Thai-French TRF-CNRS Biodiversity Project (PICS Thaïlande) and the C.N.R.S. 'Eclipse' Program.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to Y.C. (yaowalak@dmr.go.th).

The strict anaerobe *Bacteroides fragilis* grows in and benefits from nanomolar concentrations of oxygen

Anthony D. Baughn & Michael H. Malamy

Department of Molecular Biology and Microbiology, Tufts University School of Medicine, 136 Harrison Avenue, Boston, Massachusetts 02111, USA

Strict anaerobes cannot grow in the presence of greater than 5 μ M dissolved oxygen¹. Despite this growth inhibition, many strict anaerobes of the *Bacteroides* class of eubacteria can survive in oxygenated environments until the partial pressure of O₂ (P_{O_2}) is sufficiently reduced. For example, the periodontal pathogens *Porphyromonas gingivalis* and *Tannerella forsythensis* colonize subgingival plaques of mammals, whereas several other *Bacteroides* species colonize the gastrointestinal tract of animals. It has been suggested that pre-colonization of these sites by facultative anaerobes is essential for reduction of the P_{O_2} and subsequent colonization by strict anaerobes². However, this model is inconsistent with the observation that *Bacteroides fragilis* can colonize the colon in the absence of facultative anaerobes³. Thus, this

strict anaerobe may have a role in reduction of the environmental P_{O_2} . Although some strictly anaerobic bacteria can consume oxygen through an integral membrane electron transport system⁴, the physiological role of this system has not been established in these organisms. Here we demonstrate that *B. fragilis* encodes a cytochrome *bd* oxidase that is essential for O_2 consumption and is required, under some conditions, for the stimulation of growth in the presence of nanomolar concentrations of O_2 . Furthermore, our data suggest that this property is conserved in many other organisms that have been described as strict anaerobes.

To examine the effect of O_2 on the growth of *B. fragilis*, the minimum growth-inhibitory P_{O_2} was determined for the wild-type strain cultivated in a completely defined medium (anaerobic minimal medium with 0.5% glucose; AMM-gluc). When cells were shaken in sealed flasks containing a headspace with 2×10^{-3} atm O_2 (2 μ M dissolved), a slow decrease in culture viability was observed, and this decrease was more rapid with a greater concentration of O_2 (Fig. 1a). By contrast, *B. fragilis* cells grew with an initial generation time of approximately 7 h in the presence of 1×10^{-3} atm O_2 (1 μ M dissolved; Fig. 1a). After 24 h, the generation time decreased to about 3 h, similar to the anaerobic generation

time (Fig. 1a). When this change in growth rate occurred, the headspace O_2 had decreased to 5×10^{-4} atm (500 nM dissolved; data not shown). After 48 h, the headspace O_2 had decreased to less than 1×10^{-4} atm (<100 nM dissolved; data not shown).

To determine whether other members of the *Bacteroides* class can grow in the presence of low O_2 concentrations, *Bacteroides caccae*, *Bacteroides distasonis*, *Bacteroides ovatus*, *Bacteroides thetaiotaomicron*, *Bacteroides uniformis* and *Bacteroides vulgatus* were grown in AMM-gluc in a closed chamber containing 3×10^{-4} atm O_2 (300 nM dissolved). This O_2 concentration did not inhibit growth of these species (data not shown), indicating that many *Bacteroides* class members can grow in the presence of nanomolar concentrations of O_2 .

To establish whether *B. fragilis* can participate in the elimination of environmental O_2 , the rate of O_2 depletion from the culture headspace was assessed. Under the conditions tested, 10^{10} colony-forming units of wild-type *B. fragilis* in buffer consumed O_2 at a rate of 50 nmol min^{-1} (9 nmol min^{-1} per mg dry cells; Fig. 2), whereas no O_2 consumption was detected with buffer alone.

Biological reduction of O_2 typically occurs by means of terminal oxidases, such as cytochrome oxidase or NADH oxidase, and is coupled to the oxidation of reducing equivalents such as NADH. Crude cell extracts from *B. fragilis* were found to contain 5 milliunits (per mg protein) of O_2 -dependent NADH oxidase activity, indicating that this organism may couple the reduction of O_2 to the oxidation of NADH.

A two-gene cluster, *cydAB*, encoding a homologue of cytochrome *bd* oxidase was identified in the *B. fragilis* NCTC9343 and 638R preliminary genome sequences (Fig. 3). To determine whether these genes are required for O_2 depletion and O_2 -dependent NADH oxidase activity, *cydAB* was deleted from the *B. fragilis* 638R chromosome (Fig. 3). The *B. fragilis* Δ *cydAB* strain was defective for O_2 depletion relative to the wild-type strain (Fig. 2), and crude cell extracts prepared from this strain contained <0.09 milliunits (per mg protein) of O_2 -dependent NADH oxidase activity. These data indicate that *cydAB* encodes an O_2 -dependent NADH oxidase involved in O_2 consumption.

Facultative anaerobes, such as *Lactococcus lactis* and *Escherichia coli*, use cytochrome *bd* oxidase for conservation of metabolic energy and maintenance of cytoplasmic redox potential^{5,6}. When

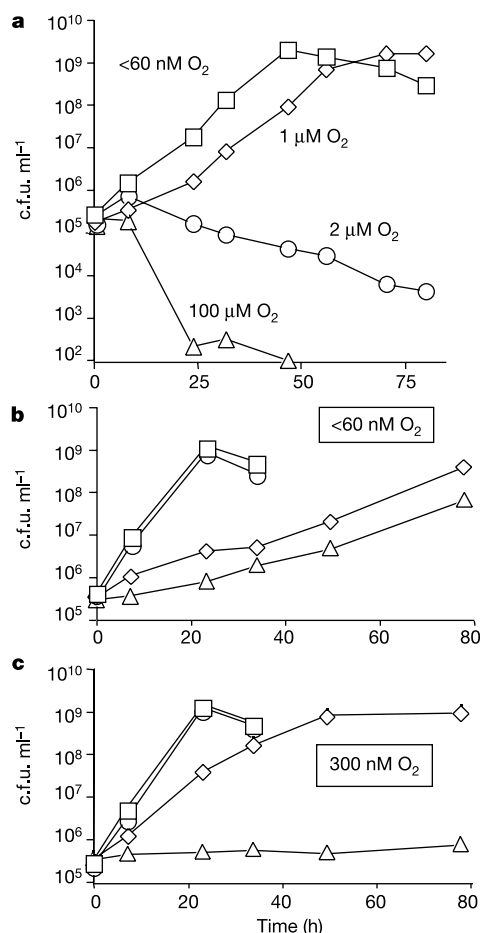


Figure 1 O_2 -dependent growth stimulation of *B. fragilis* requires cytochrome *bd* oxidase. *Bacteroides fragilis* strains were grown in an anaerobic chamber (0.85 atm N_2 , 0.1 atm H_2 , 0.05 atm CO_2 and $<6 \times 10^{-5}$ atm O_2) to mid-log phase and subcultured $1:2,000$ to AMM-gluc. **a**, *Bacteroides fragilis* ADB77 (wild type) was shaken at 100 r.p.m. in sealed flasks containing $<6 \times 10^{-5}$ atm O_2 (squares), 1×10^{-3} atm O_2 (diamonds), 2×10^{-3} atm O_2 (circles), and 0.1 atm O_2 (triangles). **b**, **c**, *Bacteroides fragilis* strains (wild type, squares; Δ *cydAB*, circles; Δ *frdC*, diamonds; Δ *cydAB* Δ *frdC*, triangles) were shaken at 300 r.p.m. in a closed 1.6×10^3 litre chamber containing $<6 \times 10^{-5}$ atm O_2 (b), or 3×10^{-4} atm O_2 (c). c.f.u., colony-forming units.

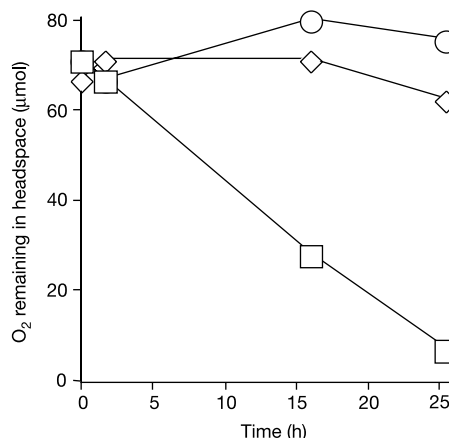


Figure 2 The *B. fragilis* *cydAB* genes are required for consumption of O_2 . *Bacteroides fragilis* strains were grown to mid-log phase as described in Fig. 1, and cells were harvested, resuspended at a density of approximately 10^8 c.f.u. ml^{-1} in 100 ml PBS containing 0.5% glucose, and placed in sealed flasks under 1 litre of 1.5×10^{-3} atm O_2 . Flasks were shaken at 300 r.p.m., $37^\circ C$. O_2 in the headspace above *B. fragilis* strains (wild type, squares; Δ *cydAB*, diamonds) and buffer alone (circles) was monitored by using the pyrogallol method²¹.

the *B. fragilis* $\Delta cydAB$ strain was cultivated in a fixed atmosphere chamber without added O_2 or with 300 nM O_2 , growth was indistinguishable from that of the wild-type strain (Fig. 1b, c). Thus, under these conditions, cytochrome *bd* oxidase is not essential for energy metabolism and redox potential maintenance in *B. fragilis*.

In *B. fragilis*, fumarate reductase is the terminal component of an anaerobic respiratory chain that couples the oxidation of NADH to the generation of ATP^{7,8}. Without added O_2 , *B. fragilis* strains lacking fumarate reductase grew with a generation time of approximately 7 h (Fig. 1b). Notably, when the *B. fragilis* $\Delta frdC$ strain was cultivated in the presence of 300 nM O_2 , the generation time decreased by nearly twofold (Fig. 1c), indicating that O_2 can serve as an efficient terminal electron acceptor in respiration. Furthermore, O_2 did not stimulate growth of the *B. fragilis* $\Delta frdC$ $\Delta cydAB$ double mutant strain, indicating that cytochrome *bd* oxidase is essential for O_2 -dependent growth stimulation (Fig. 1c). Thus, fumarate reductase and cytochrome *bd* oxidase have redundant roles in energy metabolism in the presence of 300 nM O_2 .

Unlike the *B. fragilis* $\Delta cydAB$ and $\Delta frdC$ single mutant strains, the $\Delta cydAB$ $\Delta frdC$ double mutant strain failed to grow in the presence of 300 nM O_2 (Fig. 1c). Disruption of respiratory chains frequently leads to increased production of superoxide during exposure to O_2 (refs 5, 9); thus, it is possible that the *B. fragilis* $\Delta cydAB$ $\Delta frdC$ strain cannot grow in the presence of low O_2 due to an increased superoxide burden. However, growth of this strain was not restored by addition of the superoxide-quenching compound 4-hydroxy-2,2,6,6-tetramethylpiperidine-1-oxyl (data not shown), suggesting that O_2 inhibits growth by some other mechanism.

To assess the distribution of cytochrome *bd* oxidase within eubacteria and archaea, the *B. fragilis* *CydA* sequence was compared to the microbial genome sequences at the National Centre for Biotechnology Information (NCBI) by using TBLASTN. Surprisingly, genes for *CydA* were identified in genomes of many prokaryotes that have been classified as strict anaerobes, such as *Archaeoglobus fulgidus*, *Methanosarcina barkeri*, *Geobacter sulfur-reducens*, *Desulfovibrio* species and members of the *Bacteroides* class, indicating that all of these prokaryotes may be able to grow in and benefit from the presence of low amounts of O_2 . Consistent with this prediction, *Desulfovibrio gigas* expresses an active cytochrome *bd* oxidase respiratory chain⁴. Similar to aerobic respiratory chains, this complex can couple the oxidation of NADH or succinate to the reduction of O_2 (ref. 4). Although it has been suggested that this complex participates in lowering the environmental P_{O_2} and in the maintenance of energy metabolism during exposure to O_2 (ref. 10), the physiological role of this complex has yet to be established by genetic or biochemical means.

To determine whether cytochrome *bd* oxidase is present in members of the *Bacteroides* class for which genome sequence is not available, oligonucleotide primers were designed based on highly conserved regions within the *cydA* gene of sequenced

Bacteroides class members. These primers were used in polymerase chain reactions in which the chromosomes of *B. caccae*, *B. distasonis*, *B. ovatus*, *B. thetaiotaomicron*, *B. uniformis* and *B. vulgatus* were used as templates. With the exception of *B. distasonis*, *cydA* was amplified from all *Bacteroides* class members that were tested (data not shown). Thus, the presence of cytochrome *bd* oxidase is common to this class of purportedly strict anaerobes, indicating that this respiratory complex was probably present in the *Bacteroides* class ancestor.

A multiple sequence alignment of eubacterial and archaeal *CydA* homologues was used to reconstruct a phylogenetic tree (Fig. 4). *CydA* homologues from the *Bacteroides* class were found to form a monophyletic group (Fig. 4), reinforcing the idea that cytochrome *bd* oxidase was present in an ancestor of this class. As the *Bacteroidetes* phylum is thought to have diverged from the eubacteria before the divergence of the Gram-positive bacteria and the proteobacteria¹¹, this oxidase was probably present early in the eubacterial lineage. Consistent with this model is the deep branching of the *CydA* of *Aquifex aeolicus* (Fig. 4). Phylogenetic analysis of genes for 16S ribosomal RNA indicates that the Aquificae phylum is one of the earliest diverging phyla of eubacteria for which sequence data are available^{12,13}, suggesting that cytochrome *bd* oxidase was present in some of the most ancient eubacteria. Early existence of this complex in the eubacterial lineage is consistent with the idea that there was enough O_2 to support respiration before the emergence of oxygenic photosynthesis¹⁴. Although it is the topic of ongoing debate^{15,16}, it is possible that non-biological reactions allowed early atmospheric P_{O_2} values to reach as high as 3×10^{-3} atm¹⁶.

Despite the apparent antiquity of cytochrome *bd* oxidase, it is unclear whether this oxidase was present in the last universal common ancestor. Although this complex is widely distributed in the eubacteria, the only members of the archaea currently known to encode cytochrome *bd* oxidase are within the Euryarchaeota phylum. Furthermore, no eukaryotic cytochrome *bd* oxidase homologues have been identified. These observations suggest either that

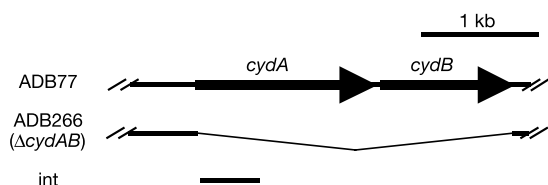


Figure 3 The *cydAB* region of *B. fragilis*. The genes encoding cytochrome *bd* oxidase (*cydAB*) are depicted by large arrows. The bar labelled ADB266 indicates the region of *cydAB* that was deleted from the *B. fragilis* TM4000 chromosome. The bar labelled 'int' represents the 0.5-kilobase (kb) fragment that was probed for and detected in other *Bacteroides* species.

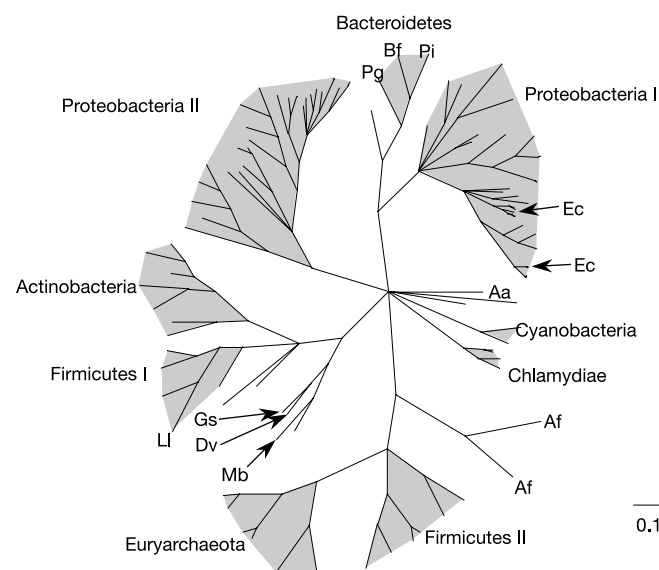


Figure 4 Phylogram of *CydA* amino acid sequences. Sequences were aligned by the CLUSTALW method. The tree was calculated by using the parsimony method with PAUP 4.0b10 (ref. 19) and visualized by using TREEVIEWPPC 1.6.5 (ref. 22). The scale represents 0.1 substitutions per residue. *CydA* sequences described in the text are marked (Aa, *A. aeolicus*; Af, *A. fulgidus*; Bf, *B. fragilis*; Dv, *Desulfovibrio vulgaris*; Ec, *E. coli*; Gs, *G. sulfurreducens*; Li, *L. lactis*; Mb, *M. barkeri*; Pg, *P. gingivalis*; Pi, *Prevotella intermedia*). Phyla from which sequences were derived are indicated. A fully annotated version of this tree is available as Supplementary Information.

this oxidase arose in the eubacteria and was subsequently acquired by an early member of the Euryarchaeota, or that this oxidase was present in the last universal common ancestor and was extensively lost in the archaeal lineage. Sequencing additional archaeal genomes and performing further phylogenetic analyses may resolve this question.

Identification of an O₂-dependent respiratory chain that can function in *B. fragilis* energy metabolism indicates that this prokaryote belongs to a new class of organisms that we term 'nanaerobes'. Similar to facultative anaerobes, nanaerobes can benefit from, yet do not require, O₂ for growth. However, similar to strict anaerobes, nanaerobes are incapable of growth in the presence of $>5 \times 10^{-3}$ atm O₂. As many other purportedly strict anaerobes encode cytochrome *bd* oxidase and grow in the presence of nanomolar concentrations of O₂, it is likely that these organisms are nanaerobes as well.

The ability of *B. fragilis* to grow in the presence of nanomolar concentrations of O₂ has important implications for the interaction of this opportunistic pathogen with its oxygenated hosts. With respect to its role as a member of the large intestinal microbiota, the ability to tolerate and use O₂ may account for the observation that *B. fragilis* is found in greatest numbers at the mucosal surface, where the P_{O₂} should be higher than it is within the intestinal lumen. During the establishment of an infection, the ability to grow in the presence of nanomolar concentrations of O₂ would enable proliferation in host tissues even before the anaerobic abscess is formed. Furthermore, the capacity to consume O₂ would allow nanaerobes such as *B. fragilis* to sustain a low local P_{O₂} in the presence of a constant influx of small amounts of O₂ from the surrounding environment. □

Methods

Strains and growth conditions

Bacteroides fragilis strains ADB77 (ref. 17), ADB247 (ref. 8), ADB266 and ADB247266 are derivatives of TM4000 (ref. 18) with the Δ thyA1 allele¹⁷. *Bacteroides caccae* ATCC43185, *B. distans* ATCC8503, *B. fragilis* ATCC23745, *B. ovatus* ATCC8483, *B. thetaiotaomicron* ATCC29741, *B. uniformis* ATCC8492 and *B. vulgatus* ATCC29327 were obtained from the Tufts Anaerobic Laboratory. *Bacteroides* species were grown in AMM-gluc¹⁷. Medium was supplemented with 50 µg ml⁻¹ thymine for growth of thy⁻ *B. fragilis* strains. Culture growth was monitored by plating for colonies on BHIS medium. All growth experiments were performed at least twice. *Bacteroides* species were grown in an anaerobic chamber ($<6 \times 10^{-5}$ atm O₂) at 37 °C with shaking at 300 r.p.m. under a gas mix of 0.05 atm CO₂, 0.10 atm H₂ and 0.85 atm N₂ in the presence of platinum–palladium catalyst to remove all traces of O₂. When necessary, 3×10^{-4} atm O₂ was included in the chamber gas mix and the platinum–palladium catalyst was removed. O₂ minimum inhibitory concentration experiments were performed using sealed flasks with Hungate side ports. Cultures were inoculated in the anaerobic chamber and flasks were sealed. O₂ was injected through the side ports using a Hamilton gas-tight syringe. Cultures were shaken at 100 r.p.m., 37 °C. Headspace-to-culture volume ratio was 40.

DNA manipulations and sequence analysis

Primary sequence data were obtained from the *B. fragilis* NCTC9343 and 638R preliminary genome sequences produced by the Pathogen Sequencing Group at the Sanger Centre (http://www.sanger.ac.uk/Projects/B_fragilis/). All other nucleotide and protein sequence data described in this study were obtained from the NCBI (<http://www.ncbi.nlm.nih.gov/>). The software programs DNA STRIDER 1.2 (DNASTar), MACVECTOR 7.0 (Oxford Molecular), PAUP 4.0b10 (ref. 19) and TOPPED 2 (ref. 20) were used for DNA and protein sequence analysis.

Bacteroides fragilis Δ cydAB strains were constructed by using the allelic exchange method as previously described¹⁷. The Δ cydAB allele was designed on the basis of sequence data obtained from the NCTC9343 preliminary genome sequence. In the Δ cydAB strains, nucleotides 44 of the *cydA* coding sequence through to 1120 of the *cydB* coding sequence were replaced with an *NcoI* site (Fig. 3).

O₂ consumption assay

Bacteroides fragilis strains were grown in AMM-gluc to an optical density at 600 nm of 0.5. Cells were harvested in sealed tubes by centrifugation at 2,300g for 5 min at room temperature. Cells were resuspended to an optical density at 600 nm of 0.1 in pre-boiled phosphate-buffered saline containing 0.5% glucose at 37 °C in the anaerobic chamber. A total of 100 ml of cell suspension was placed in a 1-litre Erlenmeyer flask with a Hungate port and the flask was sealed. By using a syringe, 7.5 ml of room air at 37 °C (0.2 atm O₂) was injected through the Hungate port to give a final concentration of 1.5×10^{-4} atm O₂. Flasks were shaken on a rotary shaker at 37 °C in the anaerobic chamber. Headspace P_{O₂} was determined by using a modification of the pyrogallol method²¹. We

withdrew 20 ml of headspace gas using a syringe. Constant pressure within flasks was maintained by injecting 20 ml of anaerobic gas mix before sampling. Gas samples were immediately injected into Hungate-type anaerobic culture tubes (Bellco Glass) containing 0.4 mM pyrogallol, 20 mM KOH. Pyrogallol solution was prepared in the anaerobic chamber with pre-boiled H₂O. Pyrogallol tubes were shaken at 300 r.p.m. on a rotary tube shaker at room temperature for 30 min. Increase in absorbance at 360 nm was determined by using a Bausch & Lomb Spectronic-20 spectrophotometer. When a 20-ml gas sample was used, the linear range of this assay was 6×10^{-5} to 2×10^{-3} atm O₂.

Enzyme assays

Unless otherwise stated, all solutions were prepared in the anaerobic chamber with pre-boiled distilled water. Enzyme assays were performed by using anaerobically prepared crude cell extracts⁸. Exponential phase cultures were washed with chilled pre-reduced buffer A (50 mM potassium phosphate pH 7.0, 5 mM MgCl₂, 0.1 M sucrose)⁷ and resuspended in 0.02 volume of buffer A. Cell suspensions were sonicated and then clarified by centrifugation for 5 min at 2,000g at 4 °C. O₂-dependent NADH oxidase activity was determined in a reaction buffer containing 50 mM potassium phosphate (pH 7.6), 5 mM MgCl₂ and 200 µM NADH in a final volume of 1 ml (ref. 8). Cell extract was added to the reaction mixture and pre-incubated for 10 min to consume any endogenous NADH oxidase substrates. Reactions were started by adding 100 µl air-saturated water (0.2 mM O₂). Cuvettes were sealed with parafilm. Activity was determined by measuring oxidation of NADH to NAD⁺ as indicated by a decrease in absorbance at 340 nm. One unit was defined as the amount of enzyme required to oxidize 1 µmol of NADH in 1 min in the presence of 20 µM O₂.

Received 27 October; accepted 12 December 2003; doi:10.1038/nature02285.

- Engelkirk, P. G., Duben-Engelkirk, J. & Dowell, V. R. *Principles and Practice of Clinical Anaerobic Bacteriology* (Star Publishing Company, Belmont, California, 1992).
- Levitt, M. D. Oxygen tension in the gut. *N. Engl. J. Med.* **282**, 1039–1040 (1970).
- Bornside, G. H., Donovan, W. E. & Myers, M. B. Intracolonic tensions of oxygen and carbon dioxide in germfree, conventional, and gnotobiotic rats. *Proc. Soc. Exp. Biol. Med.* **151**, 437–441 (1976).
- Lemos, R. S. *et al.* The 'strict' anaerobe *Desulfovibrio gigas* contains a membrane-bound oxygen-reducing respiratory chain. *FEBS Lett.* **496**, 40–43 (2001).
- Duwat, P. *et al.* Respiration capacity of the fermenting bacterium *Lactococcus lactis* and its positive effects on growth and survival. *J. Bacteriol.* **183**, 4509–4516 (2001).
- Poole, R. K. & Cook, G. M. Redundancy of aerobic respiratory chains in bacteria? Routes, reasons and regulation. *Adv. Microb. Physiol.* **43**, 165–224 (2000).
- Macy, J., Probst, I. & Gottschalk, G. Evidence for cytochrome involvement in fumarate reduction and adenosine 5'-triphosphate synthesis by *Bacteroides fragilis* grown in the presence of hemin. *J. Bacteriol.* **123**, 436–442 (1975).
- Baughn, A. D. & Malamy, M. H. The essential role of fumarate reductase in haem-dependent growth stimulation of *Bacteroides fragilis*. *Microbiology* **149**, 1503–1511 (2003).
- Kushnareva, Y., Murphy, A. N. & Andreyev, A. Complex I-mediated reactive oxygen species generation: modulation by cytochrome *c* and NAD(P)⁺ oxidation–reduction state. *Biochem. J.* **368**, 545–553 (2002).
- Cypionka, H. Oxygen respiration by *Desulfovibrio* species. *Annu. Rev. Microbiol.* **54**, 827–848 (2000).
- Weisburg, W. G., Oyaizu, Y., Oyaizu, H. & Woese, C. R. Natural relationship between *Bacteroides* and *Flavobacterium*. *J. Bacteriol.* **164**, 230–236 (1985).
- Bocchetta, M., Gribaldo, S., Sanangelantoni, A. & Cammarano, P. Phylogenetic depth of the bacterial genera *Aquifex* and *Thermotoga* inferred from analysis of ribosomal protein, elongation factor, and RNA polymerase subunit sequences. *J. Mol. Evol.* **50**, 366–380 (2000).
- Di Giulio, M. The universal ancestor was a thermophile or a hyperthermophile: Tests and further evidence. *J. Theor. Biol.* **221**, 425–436 (2003).
- Castresana, J. & Saraste, M. Evolution of energetic metabolism: the respiration-early hypothesis. *Trends Biol. Sci.* **20**, 443–448 (1995).
- Catling, D. C., Zahnle, K. J. & McKay, C. P. Biogenic methane, hydrogen escape, and the irreversible oxidation of early Earth. *Science* **293**, 839–843 (2001).
- Towe, K. M. The problematic rise of archaeal oxygen. *Science* **295**, 1419 (2002).
- Baughn, A. D. & Malamy, M. H. A mitochondrial-like aconitase in the bacterium *Bacteroides fragilis*: Implications for the evolution of the mitochondrial Krebs cycle. *Proc. Natl Acad. Sci. USA* **99**, 4662–4667 (2002).
- Tally, F. P., Snyderman, D. R., Shimell, M. J. & Malamy, M. H. Characterization of pBFTM10, a clindamycin–erythromycin resistance transfer factor from *Bacteroides fragilis*. *J. Bacteriol.* **151**, 686–691 (1982).
- Swofford, D. L. *Phylogenetic Analysis Using Parsimony and Other Methods* (Sinauer, Sunderland, 2002).
- Claros, M. G. & von Heijne, G. TopPred II: An improved software for membrane protein structure predictions. *Comput. Appl. Biosci.* **10**, 685–686 (1994).
- Schmalz, H. & Werner, H. Ein empfindlicher qualitativer Sauerstoff-Nachweis mit Pyrogallol und Kalilauge. *Berichte Deutsch. Chem. Gesellschaft* **52**, 71–73 (1925).
- Page, R. D. Treeview: An application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**, 357–358 (1996).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank A. L. Sonenshein for critical review of the manuscript, and D. W. Lazinski and A. Camilli for comments on the manuscript. This study was supported by a US Public Health Service Grant.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.H.M. (michael.malamy@tufts.edu).

A newly discovered *Roseobacter* cluster in temperate and polar oceans

Natascha Selje*, Meinhard Simon & Thorsten Brinkhoff

Institute for Chemistry and Biology of the Marine Environment,
University of Oldenburg, PO Box 2503, D-26111 Oldenburg, Germany

* Present address: Institute for Aquaculture Systems and Animal Nutrition in the Tropics and Subtropics,
University of Hohenheim, Fruwirthstrasse 12, D-70599 Stuttgart, Germany

Bacterioplankton phylotypes of α -*Proteobacteria* have been detected in various marine regions, but systematic biogeographical studies of their global distribution are missing. α -*Proteobacteria* comprise one of the largest fractions of heterotrophic marine bacteria^{1,2} and include two clades, SAR11 and *Roseobacter*, which account for 26 and 16% of 16S ribosomal RNA gene clones retrieved from marine bacterioplankton³. The SAR11 clade attracted much interest because related 16S rRNA gene clones were among the first groups of marine bacteria to be identified by cultivation-independent approaches⁴ and appear to dominate subtropical surface bacterioplankton communities⁵. Here we report on the global distribution of a newly discovered cluster affiliated to the *Roseobacter* clade, comprising only as-yet-uncultured phylotypes. Bacteria of this cluster occur from tem-

perate to polar regions with highest abundance in the Southern Ocean, but not in tropical and subtropical regions. Between the south Atlantic subtropical front and Antarctica, we detected two distinct phylotypes, one north and one south of the polar front, indicating that two adjacent but different oceanic provinces allow the persistence of distinct but closely related phylotypes. These results suggest that the global distribution of major marine bacterioplankton components is related to oceanic water masses and controlled by their environmental and biogeochemical properties.

In a study to identify abundant components of α -*Proteobacteria* in the German Bight of the North Sea, we consistently detected one phylotype affiliating with the *Roseobacter*-clade-affiliated (RCA) cluster from May 1999 to September 2001. We applied a dilution culture approach⁶ in combination with the analysis of polymerase chain reaction (PCR)-amplified bacterial 16S rRNA gene fragments in the dilution cultures and *in situ* by denaturing gradient gel electrophoresis (DGGE)⁷. The phylotype affiliating with the RCA cluster appeared regularly as a prominent band in the DGGE banding patterns of the original samples and of the highest dilution step in which growth was detected (10^{-5} or 10^{-6}). In November 1999 this phylotype was detected in all of three 10^{-6} dilutions, and in September 2001, in two of three 10^{-6} dilutions. According to the MPN index⁸, this translates into 2.3×10^5 and 1.9×10^5 cells ml⁻¹ in the original sample, corresponding to ~20% and ~7% of total DAPI cell counts. These findings indicate that the bacterium identified by this phylotype constitutes a major component of the North Sea bacterioplankton. They complement a report that,

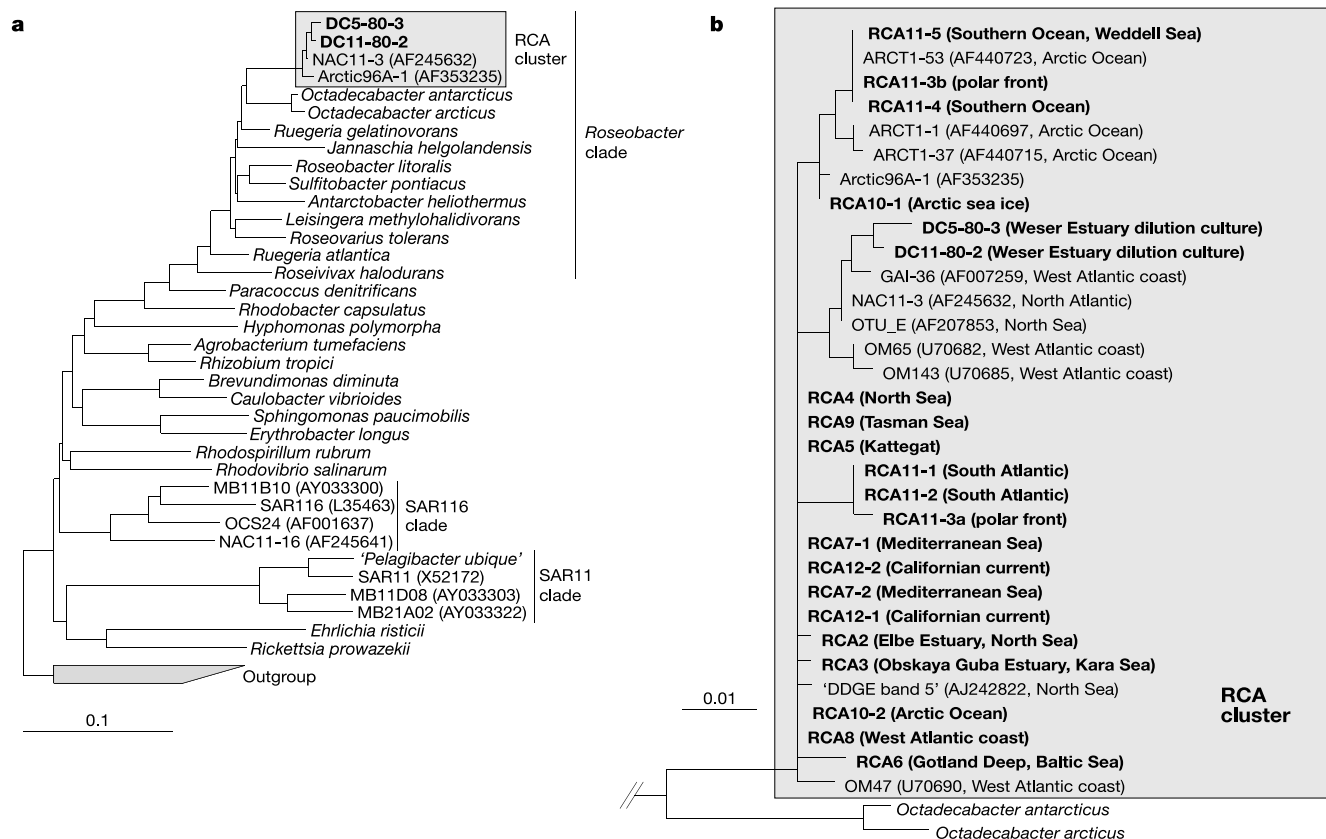


Figure 1 Phylogenetic trees showing the affiliation of the RCA cluster 16S rRNA gene sequences within α -*Proteobacteria*. **a**, Maximum-likelihood tree based on almost-complete (>1,300 nucleotides) sequences. **b**, Phylogenetic affiliation of sequences >400 bp within the RCA cluster. The tree of Fig. 1a served as a backbone tree.

Sequences <1,300 nucleotides were added with the maximum parsimony method. Note that RCA10-2 does not affiliate with RCA11-5 despite 100% identity in overlapping nucleotides, because 30 bp are missing at the end of the sequence of RCA10-2. The scale bars indicate 10% (**a**) and 1% (**b**) sequence divergence.

Table 1 Characteristics of the samples used in the global survey

Location	Latitude	Longitude	Temperature (°C)		RCA	Sampling time	N
			<i>In situ</i>	Annual range			
Arctic Ocean, Barents Sea (transect)	78°50' to 79°03' N	8°19' E to 17°32' W	−0.9–2.6	−1.9–3	+	Jul–Aug 00	24
Arctic Ocean, Kara Sea	73°15' and 74°30' N	74°0' and 74°0' E	4.3–6.3	−1.9–3	+	Aug–Oct 99	2
West Atlantic coast, Woods Hole	41°32' N	70°41' W	21	3–21	+	Aug 01	1
Caribbean Sea, Curaçao	12°10' N	69°01' W	25	25–28	−	Mar 99	3
North Sea, German Bight	53°28' to 54°22' N	6°36' to 7°30' E	6–21	1–21	+	May 99–May 01	34
German Bight, Weser Estuary (transects)	53°36' to 53°06' N	8°29' to 8°30' E	6–21	1–21	+	May 99–Nov 00	36
German Bight, Elbe Estuary (transect)	53°55' to 53°52' N	8°39' to 9°13' E	12	1–22	+	Oct 01	6
Skagerrak, Kristineberg	58°17' N	11°26' E	16	1–21	+	Sep 00	1
Kattegat, Helsingør	56°03' N	12°41' E	8	1–21	+	Apr 01	1
Baltic Sea, Gotland Deep	57°18' N	20°04' E	6–12	1–18	+	Nov 00, Oct 01	10
Baltic Sea	54°55' N	13°29' E	1–18	12	+	Oct 01	2
Mediterranean Sea, Ibiza	38°54' N	1°26' E	17	12–26	+	Nov 01	1
Mediterranean Sea, Majorca	39°33' N	2°39' E	16	12–26	+	Jan 02	1
Mediterranean Sea, Aegean Sea	35°42' N	26°31' E	20	17–26	−	Nov 01	1
Mediterranean Sea, Levantine Sea	33°37' N	29°50' E	17, 14	18–28	−	Feb 99	2
Mediterranean Sea, Levantine Sea	32°46' N	19°11' E	21	17–26	−	Nov 01	1
Central Pacific, Hawaii	19°30' N	155°55' W	24	24–27	−	Feb 02	2
Californian coast, Monterey Bay	36°75' N	122°03' W	13	12–16	+	Feb 02	2
Californian coast, Scripps Pier	32°53' N	117°15' W	15	12–22	+	Feb 02	2
Tasman Sea, Sydney	33°53' S	151°17' E	19	16–24	+	Aug 01	3
Chinese Sea, Vietnam	12°16' N	109°12' E	27	26–28	−	Sep 01	1
Red Sea, Gulf of Aqaba (transect)	27°11' to 29°29' N	34°04' to 34°57' E	20–23	20–29	−	Feb–Mar 99	12
Indian Ocean, Maldives	4°05' N	73°29' E	27	24–29	−	Oct 01	2
South Atlantic, Benguela current	25°31' S	13°33' E	18, 7	14–19	−	May 99	2
South Atlantic/Agulhas retroflection	38°57' S	19°06' E	19	17–22	−	Mar 99	1
Southern Ocean (transect)	41°20' to 70°07' S	7°44' E to 19°17' W	−1.7–4.8	−1.9–4.8	+	Mar–Apr 99	29

Locations, positions, temperature *in situ* and annual range at sea surface, presence (+) or absence (−) of the RCA cluster, sampling time, and number (N) of samples in the global survey. Water temperatures were measured *in situ* (Barents Sea, Kara Sea, German Bight, Red Sea, Southern Ocean) or derived from the Optimum Thermal Interpolation System (OTIS; <https://www.fnmoc.navy.mil/PUBLIC/OTIS/otis.html>).

during a phytoplankton bloom in the North Sea, a phylotype of another cluster of the *Roseobacter* clade with a sequence similarity of <92% to the RCA cluster was numerically abundant⁹.

Phylogenetic analysis of almost-complete 16S rRNA gene sequences from our dilution cultures in which no other organisms were detected yielded a distinct cluster within the *Roseobacter* clade with other sequences of >1,300 base pairs (bp) from the central North Atlantic¹⁰ and the Arctic oceans¹¹ (Fig. 1a). Sequences within the RCA cluster exhibit >98% similarity to and at least 4.5% difference from sequences of the next described species, *Octadecabacter antarcticus*¹². It also includes other sequences of >400 bp from temperate and polar regions^{13–15} (GenBank accession numbers AF440697, AF440715 and AF440723).

To further elucidate the occurrence of representatives of the RCA cluster, we designed a primer set for a specific PCR detection. Its application showed the presence of members of this cluster in the water column of the German Bight at various locations and over several years (1999 to 2001, Table 1), but not on solid surfaces or in freshwater habitats of adjacent estuaries.

The global distribution of the RCA cluster was investigated by screening a large set of samples by the specific PCR assay and subsequent sequencing (Table 1). This survey demonstrated the presence of the RCA cluster in the surface waters (0–40 m) of temperate to polar oceans of the Northern and Southern hemispheres (Figs 1b and 2) and to depths of 2,300 and 1,000 m in the Arctic and Southern oceans, respectively. No RCA cluster phylotypes were found in tropical and subtropical waters, including the eastern Mediterranean Sea and the south Atlantic (Benguela current, Agulhas retroflection). A global distribution of the SAR11 clade has been reported recently and time series in the subtropical Atlantic (Sargasso Sea) indicated pronounced seasonal variations in surface waters from <1% to 32% of bulk RNA with the highest proportions from June to August⁵. This suggests that temperature is one important variable controlling the distribution of the SAR11 clade and the RCA cluster and implies that both bacterial phylogenetic groups are affected differently, resulting in different global distribution patterns.

Probing a transect in the Southern Ocean between Cape Town and Antarctica revealed distinct patterns of the RCA cluster, corresponding to the oceanic provinces spanned (Fig. 2b). DGGE analysis of RCA-cluster-specific PCR products of 14 samples collected along this transect showed two bands with different migration behaviour. The RCA cluster was absent north of the subtropical front, but phylotypes affiliating with this cluster were detected in all surface samples south of the subtropical front. The sequence analysis revealed two phylotypes with 99.7% similarity (based on 912 bp; Fig. 1b). RCA11-2, occurring between the subtropical front and the polar front, and RCA11-5, occurring south of the polar front in the Antarctic circumpolar current and the Weddell Sea (Fig. 2b), differed by three nucleotides. In the polar frontal zone between 50° and 52° S, the occurrence of both phylotypes overlapped. The fragment of RCA11-5 from south of the polar front was identical to that of RCA10-2 from the Arctic Ocean (882 bp overlap).

Oceanographic biogeographical studies have shown that hydrographically separated water masses with different environmental characteristics often lead to the evolution and persistence of specific phytoplankton and zooplankton species and communities¹⁶. Our results indicate that this is also true for planktonic bacteria, as shown by the presence and absence of RCA phylotypes in the western and eastern Mediterranean, respectively, and the occurrence of different RCA phylotypes north and south of the polar front in the Southern Ocean. Obviously the different properties of the adjacent subantarctic and Antarctic water masses allowed the establishment of distinct populations of the RCA cluster, despite some mixing in the polar frontal zone. One phylotype of the RCA cluster exhibited a bipolar distribution, as has been reported for the most closely related genus, *Octadecabacter*¹⁷.

Our dilution cultures and DGGE analyses in the German Bight as well as the recurrent detection of RCA-cluster-affiliated phylotypes in the Atlantic and adjacent regions^{10,11,13–15} suggest that this cluster is a prominent bacterioplankton component in these water masses. To further substantiate the significance of the RCA cluster, we developed a new approach to estimate its fraction on the basis of

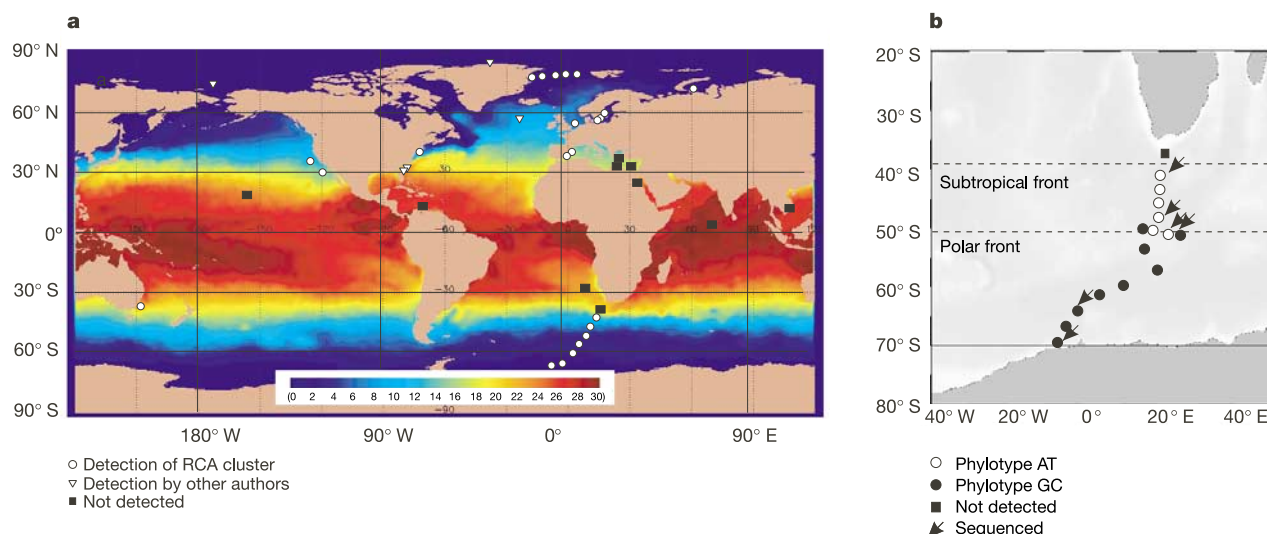


Figure 2 Distribution of RCA-cluster-affiliated phylotypes, **a**, Global investigation sites and sea surface temperatures on 21 March 2002. GenBank accession numbers of RCA phylotypes detected by other authors are given in Fig. 1b. **b**, Investigation sites along a transect from the Agulhas retroflection north of the subtropical front across the polar front in the Southern Ocean in 1999, as revealed by specific PCR, DGGE of specific PCR

serial dilutions of community DNA and their analysis by a PCR with *Bacteria*-specific primers versus one with RCA-cluster-specific primers. This comparison showed that, in the samples from the German Bight, RCA-cluster-affiliated phylotypes constituted ~10% of total *Bacteria*. Samples from the Arctic Ocean yielded ~5% and samples from the polar front and Weddell Sea in the Southern Ocean yielded ~20%. In contrast, the RCA cluster constituted <0.5% in the Baltic Sea.

To avoid overestimation, samples were subjected to this approach only if they exhibited just one distinct band in the DGGE analysis of RCA-specific PCR products and if the sequence analysis confirmed the affiliation to the RCA cluster. This approach provides an estimate of the percentage of the RCA cluster but no clear-cut quantification. The quantitative analysis of samples from the German Bight in June 2002 with a RCA-specific oligonucleotide probe using fluorescence *in situ* hybridization (FISH)¹⁸ yielded $2.0 \pm 0.96\%$ to $2.9 \pm 1.04\%$ of total microbial cells enumerated by DAPI (1.3 to $2.0 \times 10^6 \text{ ml}^{-1}$, $N = 6$), equivalent to 15.7% to 17.6% of α -*Proteobacteria* cell counts. These results, together with the MPN data from the German Bight and the observation of pronounced differences in various oceanic regions, indicate that the RCA cluster is not only widely distributed but is also abundant in temperate to polar oceans, suggesting an important role in ocean-scale biogeochemistry.

In contrast to other members of the *Roseobacter* clade that exhibit aerobic anoxygenic photosynthesis and are important in some oceanic regions^{19,20}, the RCA cluster seems to comprise only strictly heterotrophic members. Applying a PCR for *pufL* and *pufM* genes coding for the subunits of the reaction centre of bacteriochlorophyll *a* (ref. 19), we did not detect any PCR product with a dilution culture of a North Sea sample, which comprised only a member of the RCA cluster and *Marinobacter* sp., as revealed by DGGE, despite a readily amplified positive control with *Roseobacter litoralis*.

So far, global distribution patterns are known from marine ammonium oxidizers²¹, but biogeographical studies of major marine bacterioplankton groups, such as the *Cytophaga-Flavobacteria*²² and SAR11 clade^{4,5}, were based on a rather low phylogenetic resolution, even though more closely related subclusters appear to

products and subsequent sequencing. The map of the global sea surface temperature was obtained from OTIS (see Table 1). 21 March 2002 was chosen because samples in the Red Sea, the South Atlantic and Southern Ocean were collected around this season and because it represents a balanced seasonal situation.

be adapted to specific niches^{23,24}. This is the first report on the global distribution of an abundant narrow cluster of marine bacterial phylotypes. It demonstrates that one cluster with 16S rRNA gene sequence similarities of >98% can persist as an important bacterioplankton component over large oceanic areas but is restricted to temperate to polar regions. Complementing recent reports on the global distribution of the SAR11 clade⁵ and on new aerobic anoxygenic phototrophs in marine bacterioplankton, including members of the *Roseobacter* clade^{19,20}, this surprisingly wide distribution of the RCA cluster sheds new light on the structure of bacterioplankton communities and suggests that general environmental properties such as temperature are important in controlling the global distribution of abundant bacterioplankton components. It calls for more investigations of large-scale distribution patterns and the identification of key factors controlling the growth of major bacterioplankton components. □

Methods

Dilution cultures

Autoclaved sea water supplemented with vitamins and trace elements served as growth medium⁶. Incubation was done at the *in situ* temperature.

Design and application of RCA-specific primers for PCR

All DNA samples were pre-checked with *Bacteria*-specific primers 341f and 907r (ref. 7) before specific PCR. Sequences of specific primers are rca418f (5'-CCT AGG GTC GTA AAG CAC-3') and rca994r (5'-TGG TAG CAC AGG ATG-3'). Specific primers were designed using the ProbeDesign function of the ARB package (<http://www.arb-home.de/>). Specificity of primer sequences was checked with the NCBI and RDP databases (<http://www.ncbi.nlm.nih.gov/>; <http://rdp.cme.msu.edu/html/>) and resulted in at least one mismatch to other organisms for rca418f and two mismatches for rca994r. Conditions for the RCA-specific PCR were: 95°C for 2 min; 30 cycles at 95°C for 1 min, 60°C for 1 min and 72°C for 2 min; and the final elongation step at 72°C for 7 min. Conditions for the specific PCR using rca418f with a GC-clamp⁷ were: ten cycles at an annealing temperature of 62°C and 25 cycles at 61°C as the clamp influenced specificity. Specificity was checked with marine (positive controls) and freshwater (negative controls) samples from the Weser estuary and against *Roseobacter denitrificans* (strain DSM 7001), *R. algicola* (DSM 10251) and *R. gelationovorans* (DSM 5887) as further negative controls.

Samples for DNA extraction of 200 to 500 ml were concentrated on 0.2 µm Nuclepore membranes and extracted as described²⁵. Each PCR assay included a negative (without template) and a positive (German Bight sample) control. PCR products were investigated by DGGE and sequencing. At least one band of every sampling location was sequenced. To obtain more sequence information of the Arctic and Southern ocean samples, we

performed a PCR at RCA-specific conditions using the bacterial primer gm3f (ref. 26) and rca994r. Sequencing primers were gm3f, rca418f and rca994r. The affiliation of the sequences with the RCA cluster was checked by phylogenetic analysis.

Quantification of RCA phylotypes

We developed a specific PCR approach to estimate the abundance of the RCA phylotypes relative to total bacterial 16S rRNA genes based on serial dilution of extracted DNA. Alternate 1:5 and 1:2 dilution steps to extinction were applied in triplicates, to 10^7 -fold dilution or higher. PCRs with bacterial and RCA-specific primers were performed simultaneously using 60 °C for annealing. We used a PCR with primer pair 341f–907r for comparison because it amplifies a fragment consisting of approximately 560 bp (RCA-PCR 570 bp), melting temperatures, T_m , are similar (rca418f T_m = 56 °C; rca994r T_m = 54 °C; 341f T_m = 58 °C; 907r T_m = 54 °C) and the two fragments overlap in the majority of their sequences. PCR products were analysed on the same agarose gels. Gel images were edited with GelComparII (Applied Maths). Only bands differing distinctly from background noise were treated as positive results. The fraction of RCA-specific relative to total bacterial 16S rRNA genes was determined as follows. Percentage of RCA 16S rRNA genes = $\text{dil}_{\text{Bacteria}}/\text{dil}_{\text{RCA}} \times 100$, where dil_{RCA} is the highest dilution step in which RCA-specific 16S rRNA gene fragments were detected and $\text{dil}_{\text{Bacteria}}$ is the highest dilution step in which bacterial 16S rRNA gene fragments were detected. In all experiments, bacterial and RCA-cluster-specific genes were detected in all triplicates of the highest dilution steps in which they occurred.

RCA-specific oligonucleotides for FISH

Specific oligonucleotides were designed and specificity was checked as described for primer design. Owing to the low signal intensity of oligonucleotides used for the specific PCR approach, probe RCA826 was designed as 5'-ATACTTGCTGACGTCTGG-3'. In addition, probe ALF968, which targeted α -*Proteobacteria* labelled with Cy3, was applied²⁷. FISH with the RCA-specific oligonucleotide labelled with Cy3 and unlabelled helper probes RCA808-H (5'-CAITTCATCGTTACGGTG-3') and RCA845-H (5'-GGTGTGACCAACAAGT-3') was done as described¹⁸. Hybridization was done on quarters of 0.2 μm Nuclepore membranes at 46 °C for 5 h in hybridization buffer and 35% formamide. For negative controls, we used a Cy3-labelled non-EUB338 probe. The final concentration of each probe was 5 ng μl^{-1} . Washing and staining by DAPI was done essentially as described²⁸.

Phylogenetic analysis and tree construction

Obtained sequences were compared with sequences from public databases. Phylogenetic trees were constructed with the ARB software package. Sequences of >1,300 bp of at least two representative, validated type strains of every order and of representative phylotypes of the SAR116 and SAR11 clades of α -*Proteobacteria* were used, except clones SAR116 (479 bp) and SAR11 (1,107 bp). A sequence collection of γ -*Proteobacteria* served as outgroup. Alignment positions at which <50% of α -*Proteobacteria* sequences had the same residues were excluded to prevent uncertain alignments within highly variable positions.

Received 10 July; accepted 4 December 2003; doi:10.1038/nature02272.

- Hagström, A. *et al.* Use of 16S ribosomal DNA for delineation of marine bacterioplankton species. *Appl. Environ. Microbiol.* **68**, 3628–3633 (2002).
- Cottrell, M. T. & Kirchman, D. L. Community composition of marine bacterioplankton determined by 16S rRNA gene clone libraries and fluorescence *in situ* hybridization. *Appl. Environ. Microbiol.* **66**, 5116–5122 (2000).
- Giovannoni, S. & Rappé, M. in *Microbial Ecology of the Oceans* (ed. Kirchman, D. L.) 47–84 (Wiley-Liss, New York, 2000).
- Giovannoni, S. J., Britschgi, T. B., Moyer, C. L. & Field, K. G. Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**, 60–63 (1990).
- Morris, R. M. *et al.* SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**, 806–810 (2002).
- Schut, F. *et al.* Isolation of typical marine bacteria by dilution culture: growth, maintenance, and characteristics of isolates under laboratory conditions. *Appl. Environ. Microbiol.* **59**, 2150–2160 (1993).
- Muyzer, G. *et al.* in *Molecular Microbial Ecology Manual* (eds Akkermans, A. D. L., van, Elsas, J. D. & de Bruijn, F. J.) Ch. 3.4.4, 1–27 (Kluwer Academic, Dordrecht, 1998).
- American Public Health Association. *Standard Methods for Examination of Water and Wastewater Including Bottom Sediments and Sludge* 604–609 (APHA, Washington DC, 1969).
- Zubkov, M. V. *et al.* Linking the composition of bacterioplankton to rapid turnover of dissolved dimethylsulphoniopropionate in an algal bloom in the North Sea. *Environ. Microbiol.* **3**, 304–311 (2001).
- González, J. M. *et al.* Bacterial community structure associated with a dimethylsulphoniopropionate-producing North Atlantic algal bloom. *Appl. Environ. Microbiol.* **66**, 4237–4246 (2000).
- Bano, N. & Hollibaugh, J. T. Phylogenetic composition of bacterioplankton assemblages from the Arctic Ocean. *Appl. Environ. Microbiol.* **68**, 505–518 (2002).
- Gosink, J. J., Herwig, R. P. & Staley, J. T. *Octadecabacter arcticus* gen. nov., sp. nov., and *O. antarcticus*, sp. nov. nonpigmented, psychrophilic gas vacuolate bacteria from polar sea ice and water. *Syst. Appl. Microbiol.* **20**, 356–365 (1997).
- Fuchs, B. M., Zubkov, M. V., Sahm, K., Burkill, P. H. & Amann, R. Changes in community composition during dilution cultures of marine bacterioplankton as assessed by flow cytometric and molecular biological techniques. *Environ. Microbiol.* **2**, 191–201 (2000).
- Gonzalez, J. M. & Moran, M. A. Numerical dominance of a group of marine bacteria in the α -subclass of the class *Proteobacteria* in coastal seawater. *Appl. Environ. Microbiol.* **63**, 4237–4242 (1997).
- Rappé, M. S., Kemp, P. F. & Giovannoni, S. J. Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shelf off Cape Hatteras, North Carolina. *Limnol. Oceanogr.* **42**, 811–826 (1997).

- Longhurst, A. *Ecological Geography of the Sea* Ch. 10, 339–365 (Academic, San Diego, CA, 1998).
- Staley, J. T. & Gosink, J. J. Poles apart: biodiversity and biogeography of sea ice bacteria. *Annu. Rev. Microbiol.* **53**, 198–215 (1999).
- Fuchs, B. M., Glöckner, F. O., Wulf, J. & Amann, R. Unlabeled helper oligonucleotides increase the *in situ* accessibility to 16S rRNA of fluorescently labeled oligonucleotide probes. *Appl. Environ. Microbiol.* **66**, 3603–3607 (2000).
- Beja, O. *et al.* Unsuspected diversity among marine aerobic anoxygenic phototrophs. *Nature* **415**, 630–633 (2002).
- Kolber, Z. S. *et al.* Contribution of aerobic photoheterotrophic bacteria to the carbon cycle in the ocean. *Science* **292**, 2492–2495 (2001).
- Ward, B. B. & O'Mullan, G. D. Worldwide distribution of *Nitrosococcus oceanii*, a marine ammonia-oxidizing γ -proteobacterium, detected by PCR and sequencing of 16S rRNA and amoA genes. *Appl. Environ. Microbiol.* **68**, 4153–4157 (2002).
- Kirchman, D. L. The ecology of *Cytophaga-Flavobacteria* in aquatic environments. *FEMS Microbiol. Ecol.* **1317**, 91–100 (2002).
- Field, K. *et al.* Diversity and depth-specific distribution of SAR11 cluster rRNA genes from marine planktonic bacteria. *Appl. Environ. Microbiol.* **63**, 63–70 (1997).
- García-Martínez, J. & Rodríguez-Valera, F. Microdiversity of uncultured marine prokaryotes: the SAR11 cluster and the marine *Archaea* of Group I. *Mol. Ecol.* **9**, 935–948 (2000).
- Selje, N. & Simon, M. Composition and spatio-temporal dynamics of particle-associated and free-living bacterial communities in the Weser estuary, Germany. *Aquat. Microb. Ecol.* **30**, 221–236 (2003).
- Muyzer, G., Teske, A., Wirsén, C. O. & Jannasch, H. W. Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rRNA fragments. *Arch. Microbiol.* **164**, 165–172 (1995).
- Glöckner, F. O., Fuchs, B. M. & Amann, R. Bacterioplankton compositions of lakes and oceans: a first comparison based on fluorescence *in situ* hybridization. *Appl. Environ. Microbiol.* **65**, 3721–3726 (1999).
- Glöckner, F. O. *et al.* An *in situ* hybridization protocol for detection and identification of planktonic bacteria. *Syst. Appl. Microbiol.* **19**, 403–406 (1996).

Acknowledgements We thank R. Brinkmeyer, A. Bruns, B. Engelen, S. Germer, H.-P. Grossart, K. Pohlmann, U. Saint-Paul, G. Steward, M. Taylor, A. Teske and W. Zwisler for providing us with samples from various oceanic regions, A. Schlingloff for assistance in sequencing, and D. Dotschkal for FISH and *puf* gene analyses. We are grateful to D. L. Kirchman, M. A. Moran and U. Riebesell for suggestions on earlier versions of this manuscript. This work was supported by grants from the Deutsche Forschungsgemeinschaft.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to M.S. (m.simon@icbm.de). The sequences reported in this study are deposited under GenBank accession numbers AY165487–AY165505, AY145589 (DC5-80-3) and AY145625 (DC11-80-2).

Species-specific calls evoke asymmetric activity in the monkey's temporal poles

Amy Poremba¹, Megan Malloy², Richard C. Saunders², Richard E. Carson³, Peter Herscovitch³ & Mortimer Mishkin²

¹Behavioral and Cognitive Neuroscience, Department of Psychology, University of Iowa, Iowa City, Iowa 52242, USA

²Laboratory of Neuropsychology, National Institute of Mental Health, and ³PET Department, Clinical Center, National Institutes of Health, Bethesda, Maryland 20892, USA

It has often been proposed that the vocal calls of monkeys are precursors of human speech, in part because they provide critical information to other members of the species who rely on them for survival and social interactions^{1,2}. Both behavioural and lesion studies suggest that monkeys, like humans, use the auditory system of the left hemisphere preferentially to process vocalizations^{3,4}. To investigate the pattern of neural activity that might underlie this particular form of functional asymmetry in monkeys, we measured local cerebral metabolic activity while the animals listened passively to species-specific calls compared with a variety of other classes of sound. Within the superior temporal gyrus, significantly greater metabolic activity occurred

on the left side than on the right, only in the region of the temporal pole and only in response to monkey calls. This functional asymmetry was absent when these regions were separated by forebrain commissurotomy, suggesting that the perception of vocalizations elicits concurrent interhemispheric interactions that focus the auditory processing within a specialized area of one hemisphere.

Understanding where and how monkeys process auditory communicative signals could help delineate the precursor neural framework for the evolution of language. Because lateralization of language processing is a major cerebral organizing theme in humans, any similar asymmetry in monkeys could reflect an antecedent neural mechanism. Behavioural methods have indicated that macaques tend to turn the right ear to the source of species-specific vocalizations but not to the source of many other types of sound, suggesting that their left hemisphere is specialized for analysing vocalizations in particular^{5,6}. This suggestion is consistent with lesion findings demonstrating that ablation of auditory cortex on the left but not on the right impairs the animal's ability to discriminate monkey calls⁴.

We explored the physiology mediating this putative hemispheric specialization by injecting rhesus monkeys with radiolabelled 2-fluoro-2-deoxyglucose (FDG) immediately before exposing them to monkey vocalizations or other classes of sound, and then placing them in a positron emission tomography (PET) scanner to measure local cerebral metabolic activity⁷. On the basis of the results of an earlier metabolic (2-deoxyglucose) study of auditory processing in the monkey⁸, we selected five regions along the length of the superior temporal gyrus (STG) and examined each of them for evidence of asymmetrical activation (see Methods).

Significantly greater activation on the left than on the right was evoked only by species-specific monkey vocalizations and, within the STG, only at the level of the dorsal temporal pole (Fig. 1 and Supplementary Table 1). This effect was not elicited by any of the other sound classes, which included simple and complex non-vocal sounds, phase-scrambled species-specific monkey vocalizations, human speech and ambient background noise (see Methods). The

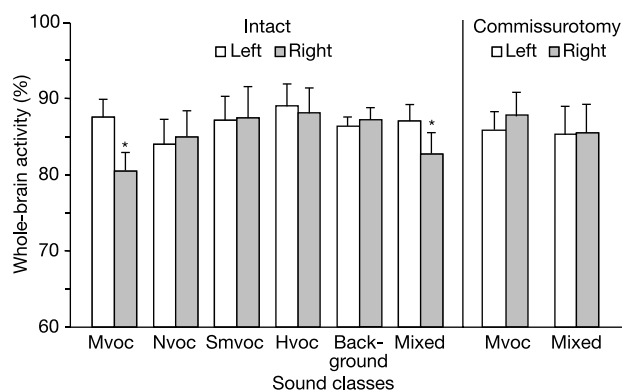


Figure 1 Metabolic activity (means and standard errors) in the monkey's dorsal temporal poles (sector 5 in the schematic brain view shown in Fig. 2) during presentation of six different classes of sound. In intact monkeys ($n = 8$), the left dorsal temporal pole was activated significantly more than the right only by species-specific monkey vocalizations, either when these were presented alone as in sound class Mvoc, or intermixed with a variety of other stimuli as in Mixed. In monkeys with forebrain commissurotomy ($n = 3$), neither of these two sound classes elicited significant asymmetric activity in the dorsal temporal pole. Activity levels were the same in both groups for the other sound classes: see Methods for descriptions of these. Asterisks indicate left activity significantly greater than right, corrected for multiple planned comparisons using Keppel's modification of the Bonferroni procedure²⁴.

effect was not only specific to monkey calls, it was sufficiently strong to remain evident even when such calls were randomly intermixed with sounds from the other classes (Fig. 1 and Supplementary Table 1).

Comparison among the activity levels evoked by the different sound classes suggested that the asymmetry observed during vocalizations could be due to trans-commissural suppression of activity in the right temporal pole. That the forebrain commissures can mediate suppression of activity in one hemisphere by activity in the other was demonstrated by, among others, behavioural studies in the motor system using both transcranial magnetic stimulation and EEG^{9,10}. To examine whether an interhemispheric suppression mechanism might have been operating in our study, we prepared monkeys with forebrain commissurotomy and then tested them in the same way we had tested the intact monkeys. No asymmetry was evident in the temporal poles of the commissurotomy cases; moreover, the activity of the right temporal pole was significantly higher in the commissurotomy than in the intact animals ($F_{1,9} = 9.295$, $P < 0.05$; Fig. 1). These findings thus provide some preliminary support for the notion that the processing of species-specific calls normally leads to trans-commissural suppression of activity in the right temporal pole.

The asymmetry in the intact animals was evident even when monkey calls were randomly intermixed with sounds from stimulus classes that did not evoke asymmetry, which implies that the relatively greater activity in the left temporal pole elicited by those calls results from a dynamic, stimulus-dependent, rather than a static, set-dependent, form of interhemispheric interaction. Stated another way, the interhemispheric interaction appears to be event-related rather than block-related, in the sense in which these terms are used in functional MRI studies¹¹, with the signal emerging from the summed activations evoked by the separated events of the same class, again as in functional MRI.

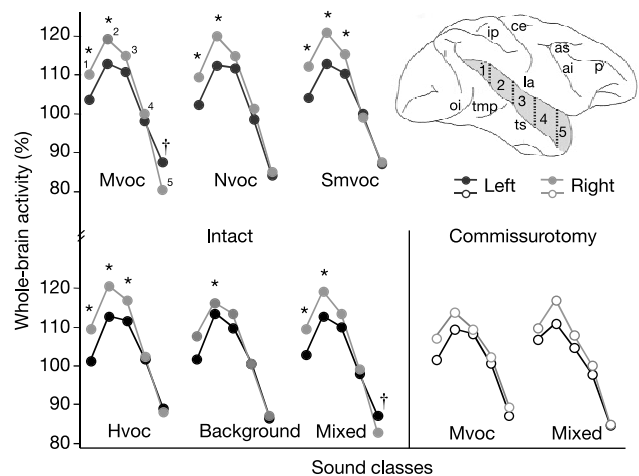


Figure 2 Metabolic activity in the left and right STG during presentation of six different sound classes (see Methods). The two rows of line graphs on the left depict activity in the intact monkeys to each of the six sound classes across STG sectors 1–5, which are demarcated on a lateral view of the monkey's brain at upper right (frontal lobe facing right). The two line graphs at the lower right depict activity in the commissurotomy monkeys to sound classes Mvoc and Mixed. The right STG is generally more active than the left across sectors 1–3 in both groups. Asterisks indicate right activity significantly greater than left, and daggers, left significantly greater than right, after correction for multiple planned comparisons using Keppel's modification of the Bonferroni procedure²⁴. Sulcal abbreviations: ai, inferior limb of arcuate; as, superior limb of arcuate; ce, central; ip, intraparietal; la, lateral; oi, inferior occipital; p, principal; tmp, posterior middle temporal; ts, superior temporal.

It is noteworthy that greater activity in the left temporal pole than in the right was observed against the background of an opposite asymmetry elsewhere in the STG. The activity levels across the STG in both hemispheres took the shape of an inverted U (Fig. 2), with the activity on both sides rising from sector 1 (the most caudal level) to reach a peak in sector 2 (the level of the primary auditory projection areas), and then decreasing progressively across sectors 3–5 (the latter at the level of the temporal pole). This inverted-U pattern in the STG is a direct reflection of cortical metabolic activity⁸. Within the inverted U, the typical pattern of asymmetry—evoked across sectors 1–3 by all sound classes, including monkey calls—was greater activity on the right than on the left (Fig. 2). This right–left difference reached significance in sectors 1 and 2 for nearly every sound class. Interestingly, the monkeys with forebrain commissurotomy showed the same pattern (Fig. 2 and Supplementary Fig. 5), although here few of the differences reached significance. Presumably this is due at least in part to the small *N*, inasmuch as the magnitude and direction of the right–left differences in the three monkeys examined were highly similar to those in the eight intact monkeys (Fig. 3, Supplementary Figs 4 and 5 and Supplementary Tables 3 and 4).

Our results suggest that within the monkey's cortical auditory system, two different types of hemispheric lateralization coexist. One type, represented in the posterior portion of STG, apparently reflects right-hemisphere specialization for processing a wide variety of acoustic stimulus classes, including at least all those presented here. This putative specialization seems to be intrinsic to the right hemisphere, in that it is largely independent of concurrent interhemispheric interaction via the forebrain commissures; whether it is also independent of these commissures ontogenetically is at present unknown. The second type of lateralization, represented in the dorsal temporal pole—a late station in the putative ventral auditory pathway⁸—apparently reflects left-hemisphere specialization for processing monkey calls specifically. This second type of lateralization depends fully on the forebrain commissures, suggesting that, in the monkey, listening to a brief call can dynamically direct cortical processing to a unilateral substrate specialized

for analysing that call. Whether the left dorsal temporal pole of the monkey is in fact necessary for analysing conspecific calls is yet to be determined.

Left-hemisphere specialization for processing of species-specific vocalizations may have an evolutionary origin in nonprimate mammals¹², paralleling that in birds¹³. Uncovering a neural basis for such specialization in monkeys, however, with their extensive auditory system⁸ and their large number of distinct vocal communicative signals^{14,15}, could yield a special benefit. A number of studies have examined the responses to vocalizations of neurons located in the monkey's primary or secondary auditory processing areas^{16–19}. Our results open up the possibility of characterizing such neuronal responses in a cortical region of the monkey that is not only a higher-order auditory processing area⁸, but also one that could be a precursor for an acoustic language area in humans. □

Methods

Subjects and stimuli

The subjects were eight unoperated rhesus monkeys (*Macaca mulatta*, 5 female, 3 male) and three operated rhesus monkeys (2 female, 1 male) with complete forebrain commissurotomy, that is, combined transection of the corpus callosum and both the anterior and hippocampal commissures; the latter were tested 3–6 months after surgery. All procedures and animal care were conducted in accordance with the National Institutes of Health (NIH) Guide for the Care and Use of Laboratory Animals. All experimental procedures were approved by the NIMH Animal Care and Use Committee.

The monkeys were placed in a sound attenuation chamber (ambient background noise ~60 dB) and presented with six classes of auditory stimuli each in a separate scanning session: (1) 'Mvoc', species-specific, unfamiliar monkey vocalizations, consisting of coos (9%), grunts (10%), barks (26%), screams (37%), gurneys/warbles (13%) and harmonic arches (5%). (2) 'Nvoc', nonvocalizations, including such environmental stimuli as glass breaking, bells ringing, water dripping, electronic sounds and so on, as well as tones, frequency sweeps, and white noise. (3) 'Smvoc', scrambled monkey vocalizations (Supplementary Fig. 6) using the same stimuli as in the first class above, but phase-scrambled in the Fourier domain, while maintaining frequency information and stimulus amplitude envelopes equal to those in the original calls. (4) 'Hvoc', human vocalizations, consisting of both male and female voices uttering words, phrases and sentences. (5) 'Background', ambient background noise consisting mainly of white noise from the building's air-handling system. (6) 'Mixed', a randomized mixture of stimuli from the above classes (excluding the phase-scrambled calls) plus music and nonprimate animal vocalizations (birds, dogs, whales and so on).

About 20% of the stimuli in the mixed condition were from the class of monkey vocalizations. The stimuli, most of which lasted from about 1 to 3 s, were presented from two speakers 26.5 cm apart positioned on the front panel of the test chamber located 24 cm in front of the monkey. During the 25 min of passive listening on each scanning day, the stimuli were presented at a rate of approximately 12 per minute with a 3-s interstimulus interval, yielding approximately 300 stimuli lasting a total of about 12 min. The order of stimulus presentations within each sound class was the same for all animals. Scanning sessions for a given animal were separated by 60 days on average.

PET scan procedure and data analysis

On the day of the scan, the monkey was seated in a primate chair and placed in a sound-attenuated chamber. An intravenous cannula (Heplock) was inserted into the animal's leg, and 15 min later it was injected with 15 mCi of FDG, which is taken up both by neurons²⁰ and by glial cells involved in neurotransmitter metabolism²¹, thereby providing a signal of energy consumption that increases linearly with functional activation. The acoustic stimuli were then presented, and after 25 min of passive listening, the monkey was sedated with ketamine (10 mg per kg intravenously) and a combined ketamine (8 mg per kg)/xylazine (0.4 mg per kg intramuscularly) solution and then transported to the PET scanner (GE Advance) for data acquisition. The animal's head was positioned within a stereotaxic head holder so as to obtain coronal images of FDG uptake.

Images were acquired in two-dimensional mode at two interleaved axial levels (2 mm apart) to provide maximal sampling. An 8-min transmission scan for attenuation correction at axial level 1 was followed by eight 5-min emission scans at alternating axial levels, followed by an 8-min transmission scan at level 2. This interleaved scan mode ensured that the average acquisition times postinjection at the two levels were nearly identical. The four image frames at each level were summed, and the slices from the two levels were interleaved to produce a final image volume of 70 slices with voxel dimension of 1 × 1 × 2 mm. The data were analysed using MedX 4.1 software. PET images were co-registered with each monkey's MRI scan using automated image registration²². MRI was performed in a 1.5-T Signa unit (GE Medical Systems), using a 5-inch surface coil. T1-weighted magnetic resonance images were obtained using three-dimensional volume SPGR pulse sequence (TE 6, TR 25, flip angle 30), field of view 11 cm, and slice thickness 1 mm. After co-registration, a normalized activity image was created by dividing the radioactivity of each voxel by the average concentration of whole-brain radioactivity. Owing to the linearity of the FDG model, these normalized values are nearly equal to normalized metabolic rates. (Normalized radioactivity values were used as the outcome

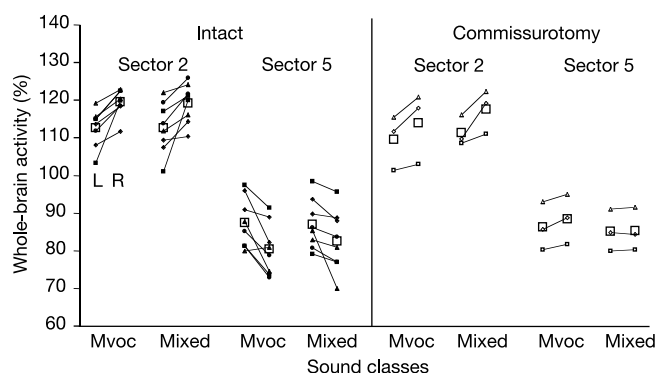


Figure 3 Metabolic activity values in sectors 2 and 5 for each individual subject for each of two sound classes, Mvoc and Mixed. Individual subject means for the left and right hemispheres are marked by the paired left and right (L and R) symbols, respectively, connected by solid lines. Group means for the left and right hemispheres are represented by the unconnected outline squares on the left and right, respectively. Sectors are demarcated on the lateral view of the monkey's brain in Fig. 2. Graphs indicate a consistent pattern of right activity greater than left across individual subjects in sector 2 (encompassing the primary auditory projection areas) for both sound classes and both groups. An equally consistent reversal of lateralization across individual subjects, with left hemisphere activity greater than right, is evident in sector 5 (temporal pole) for the same two sound classes, but for the intact monkeys only. The commissurotomy cases show no lateralization in sector 5.

measure rather than absolute values of glucose utilization rates in order to avoid the effects of blood sampling during the scanning session.)

The regions of interest (ROIs), drawn on the coronal MRI scans for each monkey, were five sectors (1–5) along the lengths of the left and right STG. According to the architectonic schema of ref. 23, sector 1, the posterior auditory cortex, includes mainly the temporoparietal cortex, caudal parakoniocortex and retroinsular temporal area; sector 2, containing the primary auditory cortex, includes mainly lateral parakoniocortex, auditory koniocortex, lateral portion of temporalis superior (Ts) 3 and prokoniocortex; sector 3 includes the major portion of Ts3 and rostral parakoniocortex; sector 4 includes the major portion of (Ts2) and anterior portions of Ts3; and sector 5, the temporal pole, includes the major portion of (Ts1) and a small dorsal portion of proisocortex. These template ROIs were applied to each realigned and normalized PET scan.

The five sectors were preselected on the basis of their differential levels of 2-deoxyglucose uptake observed during an acute metabolic mapping study of auditory processing in monkeys⁸. In line with one of the premises of ROI analyses, each sector was treated as being independent of the others. Percentages of whole-brain activity for each sector were averaged across multiple coronal sections on which that ROI appeared, and comparisons were made between hemispheres with two-tailed paired *t*-tests. Keppel's modification of the Bonferroni procedure was used to correct the 'family-wise' error rate for multiple planned comparisons²⁴, as follows. The six different stimulus classes were entered as the experimental treatment, the number of degrees of freedom for this treatment source of variance ($6 - 1 = 5$) was multiplied by the standard critical probability level (0.05), and the product was divided by the number of planned comparisons (6), yielding a new, corrected, critical probability level of 0.042. In the case of the between-group comparison within sector 5, a three-way repeated-measures analysis of variance (ANOVA) was used with group (intact and commissurotomy), stimulus class (Mvoc and Mixed) and hemispheric side as the factors.

Received 24 July; accepted 5 December 2003; doi:10.1038/nature02268.

- Cheney, D. L. & Seyfarth, R. M. *How Monkeys See The World* (Univ. Chicago Press, Chicago, IL, 1990).
- Prell, C. G., Hauser, M. D. & Moody, D. B. Discrete or graded variation within rhesus monkey screams? Psychophysical experiment on classification. *Anim. Behav.* **63**, 47–62 (2002).
- Ghazanfar, A. A. & Hauser, M. D. The auditory behaviour of primates: a neuroethological perspective. *Curr. Opin. Neurobiol.* **11**, 712–720 (2001).
- Heffner, H. E. & Heffner, R. S. Temporal lobe lesions and perception of species-specific vocalizations by macaques. *Science* **226**, 75–76 (1984).
- Petersen, M. R., Beecher, M. D., Zoloth, S. R., Moody, D. B. & Stebbins, W. C. Neural lateralization of species-specific vocalizations by Japanese Macaques. *Science* **202**, 324–327 (1978).
- Hauser, M. D. & Anderson, K. Left hemisphere dominance for processing vocalizations in adult, but not infant, rhesus monkeys: Field experiments. *Proc. Natl Acad. Sci. USA* **91**, 3946–3948 (1994).
- Phelps, M. E. *et al.* Tomographic measurement of local cerebral glucose metabolic rate in humans with (F-18) 2-fluoro-2-deoxy-D-glucose: Validation of method. *Ann. Neurol.* **6**, 371–388 (1979).
- Poremba, A. *et al.* Functional mapping of the primate auditory system. *Science* **299**, 568–572 (2003).
- Ferbert, A. *et al.* Interhemispheric inhibition of the human motor cortex. *J. Physiol. (Lond.)* **453**, 525–546 (1992).
- Netz, J. in *Transcranial Magnetic Stimulation* (eds Paulus, E., Hallett, M., Rossini, P. M. & Rothwell, J. C.) Ch. 14 (Elsevier Science, New York, 1999).
- Otten, L. J., Henson, R. N. & Rugg, M. D. Sate-related and item-related neural correlates of successful memory encoding. *Nature Neurosci.* **5**, 1339–1344 (2002).
- Ehret, G. Left hemisphere advantage in the mouse brain for recognizing ultrasonic communication calls. *Nature* **325**, 249–251 (1987).
- George, I., Cousillas, H., Richard, J. P. & Hausberger, M. Song perception in the European starling: hemispheric specialization and individual variations. *C. R. Biol.* **325**, 197–204 (2002).
- Snowdon, C. T., Brown, C. H. & Petersen, M. R. *Primate Communication* (Cambridge Univ. Press, Cambridge, UK, 1982).
- Hauser, M. D. Functional referents and acoustic similarity: field playback experiments with rhesus monkeys. *Anim. Behav.* **55**, 1647–1658 (1998).
- Glass, I. & Wollberg, Z. Responses to cells in the auditory cortex of awake squirrel monkeys to normal and reversed species-specific vocalizations. *Hear. Res.* **9**, 27–33 (1983).
- Wang, X., Merzenich, M. M., Beitel, R. & Schreiner, C. E. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. *J. Neurophysiol.* **74**, 2685–2706 (1995).
- Rauschecker, J. P. & Tian, B. Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc. Natl Acad. Sci. USA* **97**, 11800–11806 (2000).
- Wang, X. & Kadia, S. C. Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. *J. Neurosci.* **86**, 2616–2620 (2001).
- Sokoloff, L., Takahashi, S., Gotoh, J., Driscoll, B. F. & Law, M. J. Contribution of astroglia to functionally activated energy metabolism. *Dev. Neurosci.* **18**, 344–352 (1996).
- Magistretti, P. Cellular bases of functional brain imaging: insights from neuron-glia metabolic coupling. *Brain Res.* **886**, 108–112 (2000).
- Woods, R., Mazzotta, J. & Cherry, S. MRI-PET registration with automated algorithm. *J. Comput. Assist. Tomogr.* **17**, 536–546 (1993).
- Yeterian, E. H. & Pandya, D. N. Corticostriatal connections of the superior temporal region in rhesus monkeys. *J. Comp. Neurol.* **399**, 384–402 (1998).
- Keppel, G. *Design and Analysis: A Researcher's Handbook*. 2nd edn, Ch. 8 (Prentice-Hall, Englewood Cliffs, NJ, 1982).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank E. Moreton for programming a software package to phase-scramble the sounds. This work was supported by the Intramural Research Program of NIMH, NIH, DHHS, and by internal funds from the University of Iowa.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.P. (amy-poremba@uiowa.edu).

SOL-1 is a CUB-domain protein required for GLR-1 glutamate receptor function in *C. elegans*

Yi Zheng, Jerry E. Mellem, Penelope J. Brockie, David M. Madsen & Andres V. Maricq

Department of Biology, University of Utah, Salt Lake City, Utah 84112-0840, USA

Ionotropic glutamate receptors (iGluRs) mediate most excitatory synaptic signalling between neurons. Binding of the neurotransmitter glutamate causes a conformational change in these receptors that gates open a transmembrane pore through which ions can pass. The gating of iGluRs is crucially dependent on a conserved amino acid that was first identified in the 'lurcher' ataxic mouse¹. Through a screen for modifiers of iGluR function in a transgenic strain of *Caenorhabditis elegans* expressing a GLR-1 subunit containing the *lurcher* mutation, we identify suppressor of *lurcher* (*sol-1*). This gene encodes a transmembrane protein that is predicted to contain four extracellular β -barrel-forming domains known as CUB domains^{2,3}. SOL-1 and GLR-1 are colocalized at the cell surface and can be co-immunoprecipitated. By recording from neurons expressing GLR-1, we show that SOL-1 is an accessory protein that is selectively required for glutamate-gated currents. We propose that SOL-1 participates in the gating of non-NMDA (N-methyl-D-aspartate) iGluRs, thereby providing a previously unknown mechanism of regulation for this important class of neurotransmitter receptor.

In most nervous systems, fast glutamatergic neurotransmission is mediated by iGluRs that are classified on the basis of their response to the specific ligands NMDA (NMDA receptors), and AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazole propionic acid) and kainate (non-NMDA receptors)⁴. Although many molecules that interact with iGluRs have been described⁵, it is likely that additional molecules important for the function of iGluRs remain to be identified. To address this, we undertook a systematic genetic approach to identify genes that are important for non-NMDA iGluR function in the soil nematode *C. elegans*.

Our strategy was based on a gain-of-function mutation found in the *lurcher* mutant mouse¹. The mutation is located near the receptor pore region of the GluR δ 2 subunit and results in the exchange of a highly conserved alanine for threonine. By an unknown mechanism, this mutation results in both constitutive channel activation and modified gating kinetics⁶. When the homologous mutation is introduced into the *C. elegans* non-NMDA-type GLR-1 subunit^{7,8}, transgenic nematode worms expressing the modified subunit GLR-1(A/T), referred to here as '*lurcher*' worms, show a marked hyper-reversal behaviour⁹. We used this sensitized genetic background in a screen for suppressors of the *lurcher* movement phenotype. When tested in an assay that measured the time required for a population of worms to reach a

food target, nearly all of the wild-type worms and *glr-1* mutants reached the target within 2 h, whereas *lurcher* worms took far longer (Fig. 1a). Thus, we reasoned that genetic suppressors of *lurcher* would identify gene products important for iGluR expression, membrane insertion or function.

We mutagenized *lurcher* worms (see Methods) and screened for the rare mutants that reached the food within 1 h. We isolated those suppressors that restored the movement of transgenic *lurcher* worms to near that of wild-type worms (Fig. 1b). Of the mutants identified, three strong suppressors (*ak63*, *ak66* and *ak67*) fell into one complementation group that defined the *sol-1* gene. We mapped the *sol-1* locus to linkage group 1 between the markers *bli-4* and *dpy-24* (Supplementary Fig. S1). Using a technique of transformation rescue, we determined that cosmid C15A11 and a 13.7-kilobase (kb) fragment of this cosmid containing a predicted coding sequence (C15A11.3; ref. 10) were sufficient to restore the hyper-reversal behaviour of transgenic *sol-1(ak63); lurcher* worms (data not shown). We identified a full-length complementary DNA that has a SL1 *trans*-splice leader sequence and a poly(A) tail, indicating that the whole cDNA had been identified (see Methods). For each of the *sol-1* alleles, we identified a single nucleotide change in the predicted genomic coding region. *ak63* modifies an intron donor splicing site, causing premature truncation of the protein;

ak67 is a nonsense mutation, also predicted to result in a truncated protein; and *ak66* is a missense mutation, predicted to cause a single amino acid change (G323R; Fig. 1c).

The cDNA encodes a predicted protein of 594 amino acids, including a signal sequence, a transmembrane domain in the carboxy-terminal region and potential glycosylation sites (Fig. 1c). These features suggest that *sol-1* is likely to encode a type I transmembrane protein that is almost completely extracellular. Domain analysis detected the presence of four putative extracellular CUB domains (Methods). The CUB domain is a fragment of about 110 residues that was first identified² in complement subcomponents Clr/Cls, epidermal growth factor related sea urchin protein, and bone morphogenetic protein 1 (BMP1). The CUB domain is predicted to form a β -barrel similar to that found in immunoglobulins³. (An alignment of CUB domains is shown in Supplementary Fig. S2.) CUB domains have been found in the extracellular region of many functionally diverse proteins, ranging from neuropilin (a semaphorin co-receptor) to epithelial receptors such as cubulin, metalloproteases such as BMP1 and tolloid, and adhesion molecules such as spermadhesin^{2,3,11}.

As *sol-1* is required for GLR-1(A/T) function, we examined whether *sol-1* is also required for the function of wild-type GLR-1. We separated *sol-1* from the GLR-1(A/T) transgene and compared

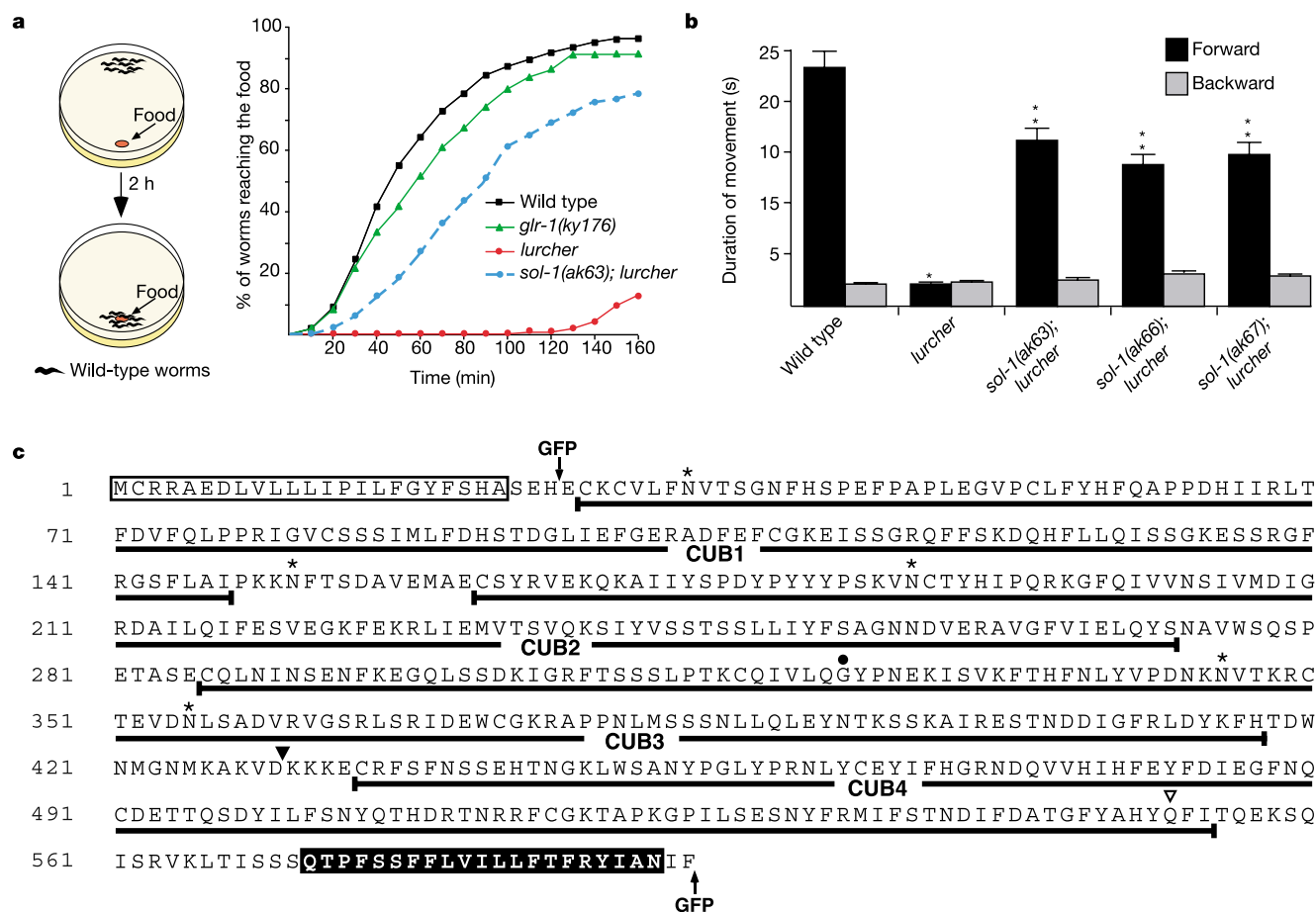


Figure 1 Genetic screen for mutations suppressing the hyper-reversal phenotype of *lurcher* worms. **a**, Worms were placed at one pole of a 10-cm agar plate (left) and allowed to migrate towards a food source at the opposite pole for 2 h. The cumulative distribution of worm arrival times is shown on the right. **b**, Mean duration of forward and backward movement of wild-type, *lurcher* and *sol-1; lurcher* mutants. Asterisks indicate a statistical difference from wild type (*) or both wild type and *lurcher* (**), $P < 0.001$. For each strain, $n = 10$. **c**, Predicted amino acid sequence encoded by *sol-1*. Amino acids are numbered

from the first predicted methionine. Shown are the putative signal sequence (open box), CUB domains (underlined) and transmembrane domain (filled box). The filled arrowhead, filled circle and open arrowhead represent sites affected by the *ak63*, *ak66* and *ak67* mutations, respectively. Asterisks indicate putative *N*-linked glycosylation sites. The positions of the N-terminal (GFP::SOL-1) and C-terminal (SOL-1::GFP) chimaeras are also indicated.

the behavioural defects of *sol-1* and *glr-1* mutants (see Methods). *glr-1* mutants have defects in tactile avoidance, duration of forward movement and osmotic avoidance^{7,8,12}. If *sol-1* is important for GLR-1 expression, localization or function, then *sol-1* and *glr-1* mutants should have similar defects. Indeed, both mutants showed similar defects in the nose touch assay and reversal frequency (duration of forward movement; Fig. 2a, b). *glr-1* and *sol-1* mutants were not defective in classical assays of osmotic avoidance behaviour (Fig. 2c), but did show delayed reaction times when acutely exposed to a drop of hyperosmotic solution (refs 12, 13, and Fig. 2d).

For each behaviour, we showed that *sol-1* mutants were rescued by transgenes expressing either wild-type SOL-1 or a functional protein chimaera in which green fluorescent protein (GFP) was fused to the amino terminus of SOL-1 (GFP::SOL-1) (Figs. 2a, b, d). All three alleles of *sol-1* are recessive and resulted in similar behavioural defects (data not shown). Because no deficiencies uncovered *sol-1* (Supplementary Fig. S1), however, we could not determine whether these mutations represent true null alleles.

To address the possibility that *sol-1* is required for GLR-1 expression, we examined the expression of a functional GLR-1::GFP chimaera¹⁴ in transgenic *sol-1* mutants. We found no obvious differences in GLR-1::GFP expression between wild-type and *sol-1* mutants (Fig. 3a–d). Notably, in *sol-1* mutants GLR-1::GFP expression was still observed in the command interneurons and as ventral cord puncta that probably represent points of synaptic contact¹⁴. Although GLR-1 seemed to be present at synapses in *sol-1* mutants, it may be that the receptor subunit is subcellular and thus nonfunctional. Therefore, we examined GLR-1 surface expression using a technique for growing dissociated neurons in culture¹⁵. GLR-1::GFP fluorescence was observed in both the cell body and processes of neurons cultured from either transgenic wild-type or *sol-1* mutants (Fig. 3e, i). Under non-permeabilized conditions, an antibody to an extracellular epitope of GLR-1 labelled only cell-surface GLR-1::GFP (Fig. 3f, g, j, k), whereas under

conditions that promote membrane permeabilization, both cell-surface and intracellular GLR-1::GFP were labelled (Fig. 3h, l). Thus, it seems that SOL-1 is not required for surface expression of GLR-1.

The interpretation of these data is confounded by the possibility that the properties of dissociated neurons are altered as compared with neurons *in vivo*. We therefore injected fluorescently labelled anti-GFP antibodies into transgenic worms expressing variants of GFP-tagged GLR-1 (see Methods). For both wild-type and *sol-1* mutants, we observed intense antibody labelling of neuronal processes in transgenic worms expressing a functional protein chimaera in which GFP was fused to the extracellular N terminus of GLR-1 (Fig. 3m, n). However, antibody labelling was absent in transgenic wild-type worms expressing GLR-1::GFP (ref. 14 and Fig. 3o; GFP is intracellular). Thus, we conclude that SOL-1 is not required for cell-surface expression of GLR-1.

Most of the glutamate-gated current in AVA interneurons activates and desensitizes rapidly, is dependent on GLR-1 and can be activated by kainate (ref. 12 and Fig. 4a, c). By contrast, NMDA-activated currents show slower kinetics, as well as voltage dependence⁴. In *sol-1* mutants, the fast glutamate-gated currents were found to be largely abolished and the kainate-gated currents were eliminated (Fig. 4b). Conversely, the currents elicited in response to NMDA were indistinguishable from those in wild-type, indicating that loss of *sol-1* function does not simply cause a nonspecific reduction in ligand-gated currents. Acetylcholine-gated currents also can be recorded from AVA interneurons; these currents were not affected by *sol-1* mutations (data not shown), providing further evidence for the specificity of SOL-1 action. Notably, the change in glutamate-gated currents caused by *sol-1* mutations exactly mimicked the change observed in *glr-1* mutants (Fig. 4c). Normal currents were restored in transgenic *sol-1* mutants expressing a *sol-1* genomic clone (Fig. 4d). To determine whether SOL-1 expression is required in AVA interneurons, we used the *nmr-1* promoter to express GFP::SOL-1 in transgenic *sol-1* mutants^{16,17}. Again, we

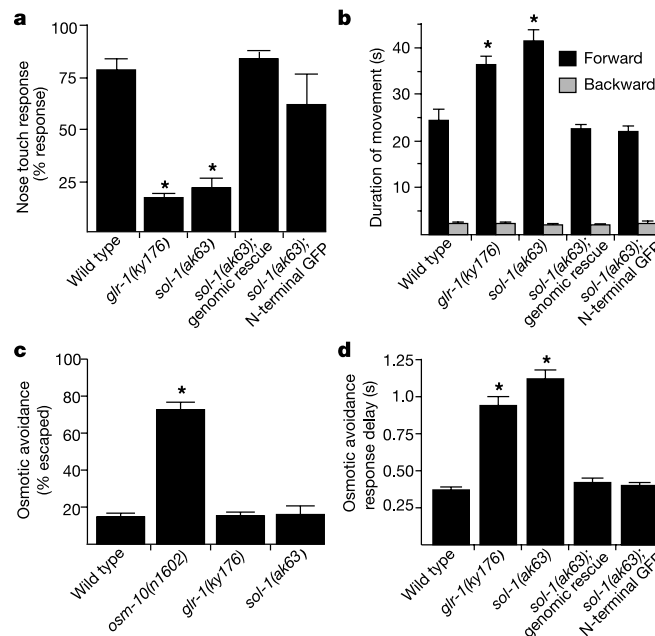


Figure 2 *sol-1(ak63)* phenocopies *glr-1(ky176)*. **a**, Percentage response to ten consecutive tactile stimuli applied to the tip of the worm's nose. **b**, Mean duration of forward and backward movement. **c**, Percentage of worms that escaped a ring of hyperosmotic fructose in 20 min (chronic exposure). For each strain in **a–c**, $n = 10$. **d**, Delay in response to a drop of fructose placed in the worm's path (acute exposure).

Genomic rescue strain: *sol-1(ak63); lin-15(n765ts); akEx204*. N-terminal GFP strain: *sol-1(ak63); lin-15(n765ts); akEx241*. Asterisk indicates a statistical difference from wild type, $P < 0.001$. Wild type, $n = 28$; *ky176*, $n = 21$; *ak63*, $n = 25$; *ak63*; genomic rescue, $n = 21$; *ak63*; N-terminal GFP, $n = 17$.

observed full rescue of the glutamate-gated current (Fig. 4d), indicating that cell-autonomous SOL-1 expression is required for GLR-1-dependent currents.

Our behavioural and electrophysiological data indicate that SOL-1 is required for GLR-1 function. To determine whether SOL-1 and GLR-1 are coexpressed, we examined transgenic worms expressing a functional chimera of yellow fluorescent protein (YFP) and SOL-1 using the *sol-1* native promoter, together with cyan fluorescent protein (CFP) driven by the *glr-1* promoter. YFP::SOL-1 was widely distributed in the nervous system, including a large subset of *glr-1*-expressing neurons (Fig. 5a). However, SOL-1 was also expressed in additional neurons that apparently did not express GLR-1. When we repeated this experiment using the *nmr-1* promoter to drive CFP^{16,17}, SOL-1 was expressed in the command interneurons, including AVA (Fig. 5b). As GLR-1 is known to be

expressed in the command interneurons^{7,8}, we conclude that SOL-1 is coexpressed with GLR-1 in these neurons.

To determine whether SOL-1 and GLR-1 colocalize in the processes of the command interneurons, we used the *nmr-1* promoter to coexpress functional GLR-1::CFP and SOL-1::YFP in transgenic worms. Discrete, overlapping puncta were observed in the neuronal processes (Fig. 5c), indicating that SOL-1 and GLR-1 are closely colocalized in the postsynaptic membrane. Do SOL-1 and GLR-1 physically associate? To address this possibility, we cotransfected COS-7 cells with functional epitope-tagged variants of SOL-1 (HA::SOL-1) and GLR-1 (GFP::GLR-1). Co-immunoprecipitation studies showed that antibodies against haemagglutinin A (HA) that precipitated HA::SOL-1 also co-precipitated GFP::GLR-1, but not a control membrane-associated variant¹⁸ of CFP (data not shown) or full-length GFP::NMR-1 (Fig. 5d). Likewise, antibodies

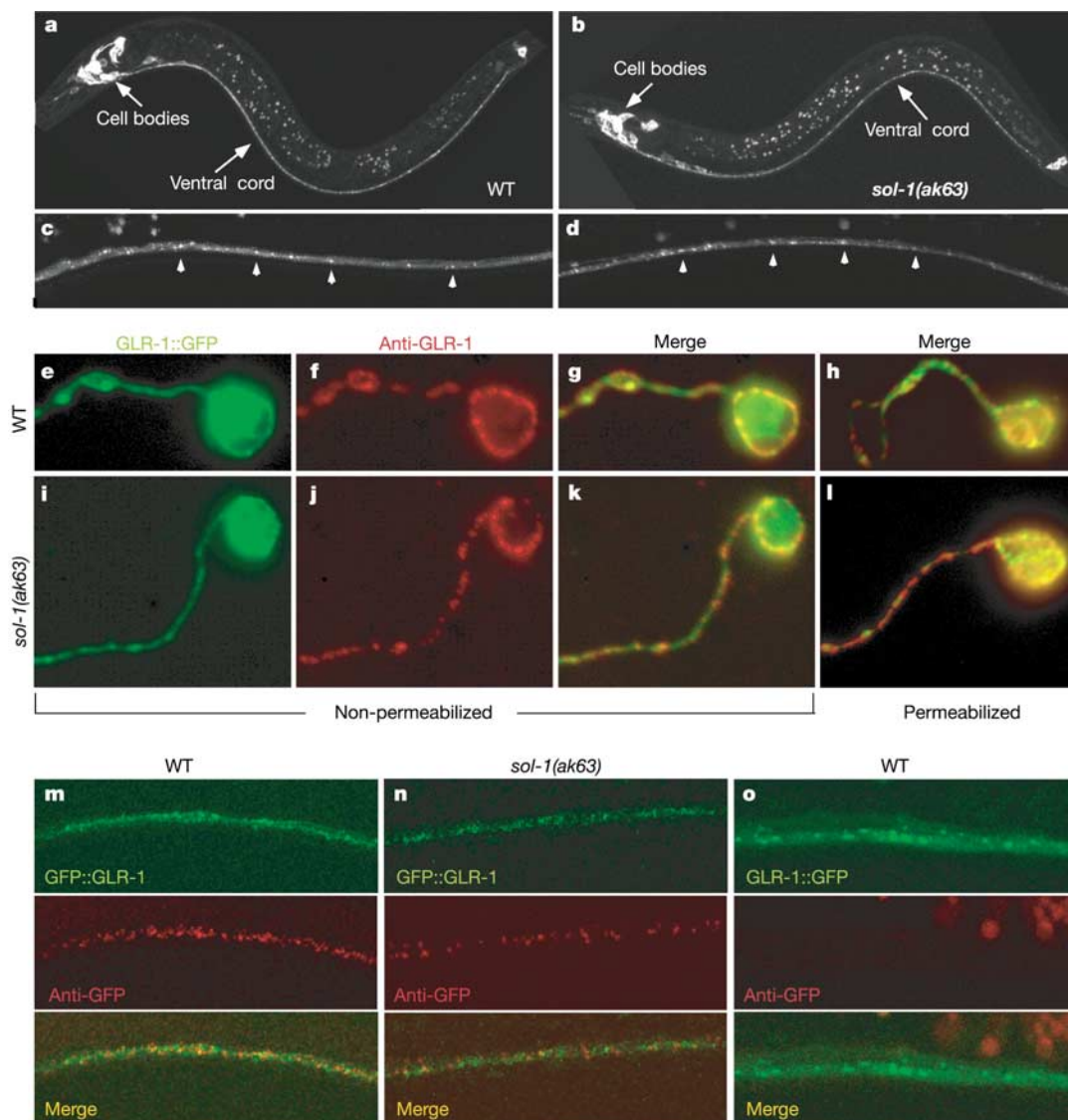


Figure 3 *sol-1(ak63)* does not affect GLR-1 expression, localization or membrane insertion. **a–d**, Confocal images of transgenic wild-type (**a, c**) and *sol-1(ak63)* (**b, d**) worms expressing GLR-1::GFP¹⁴. **a, b**, L2 larva oriented with anterior regions to the left and ventral side down (autofluorescence of fat granules in lateral positions is also apparent). **c, d**, GFP puncta (arrowheads) in the ventral neural processes of adult transgenic worms. **e–h**, GLR-1::GFP expressing neurons cultured from either wild-type (**e–h**) or *sol-1(ak63)* (**i–l**) transgenic worms. Shown are GLR-1::GFP (**e, i**), anti-GLR-1

staining (**f, j**) and the merged images (**g, k**) in non-permeabilized cells, and the merged images of GLR-1::GFP and anti-GLR-1 staining in permeabilized cells (**h, l**). See Methods for details. **m–o**, Confocal images of GFP (top), anti-GFP staining (middle) and the merged images (bottom) in transgenic wild-type (**m, o**) and *sol-1(ak63)* (**n**) worms expressing GFP::GLR-1 (extracellular GFP, **m, n**) or GLR-1::GFP (intracellular GFP, **o**) injected with Alexa594-conjugated rabbit anti-GFP polyclonal sera.

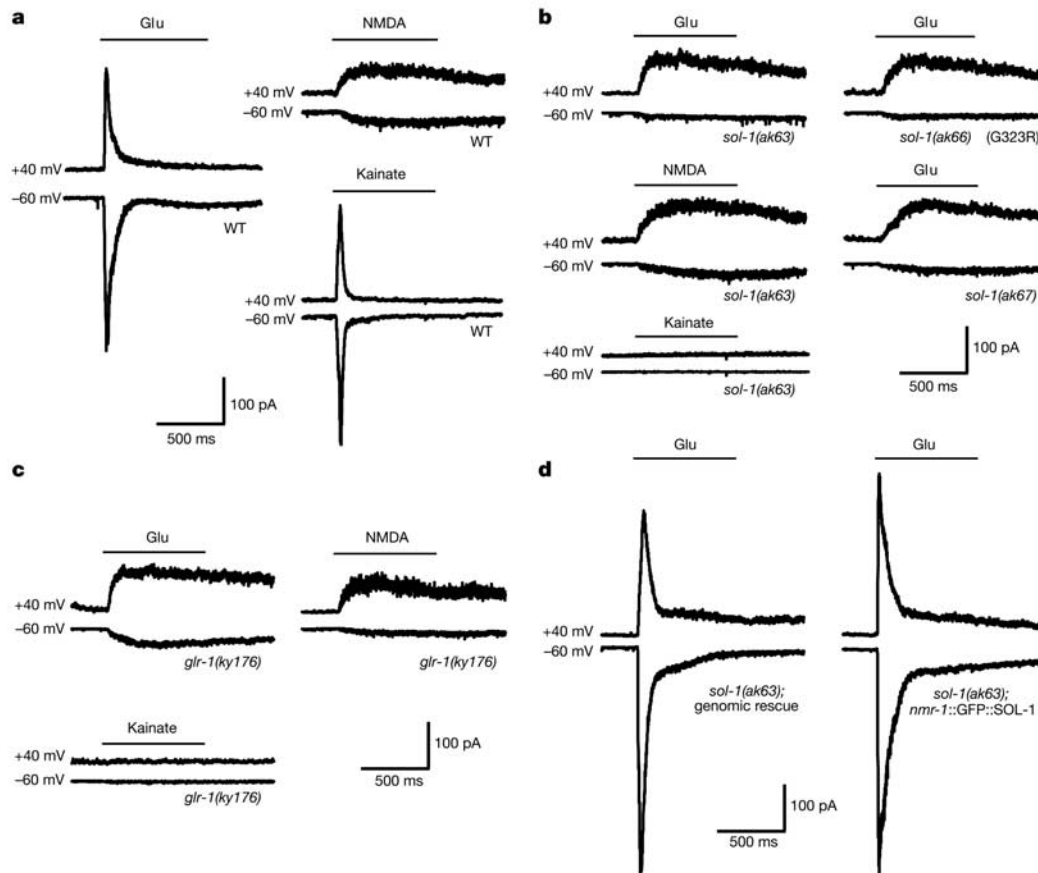


Figure 4 Non-NMDA-gated currents are eliminated in both *glr-1(ky176)* and *sol-1(ak63)* mutants. **a–c**, Currents activated by glutamate (1 mM), kainate (1 mM) or NMDA (1 mM) in the AVA interneurons of wild-type (**a**), *sol-1(ak63)* (**b**) and *glr-1(ky176)* (**c**) worms. *sol-1(ak66)* and *sol-1(ak67)* also disrupted fast glutamate-gated currents (**b**). **d**, Glutamate-gated currents were fully restored in *sol-1(ak63)* transgenic worms

expressing either a wild-type *sol-1* transgene (left) or a GFP::SOL-1 chimera driven by the *nmr-1* promoter (right). To detect the non-NMDA-dependent currents more easily, we omitted spermine from the pipette solution; therefore, these currents did not show spermine-mediated inward rectification¹².

against GFP precipitated both GFP::GLR-1 and GFP::NMR-1, but HA::SOL-1 only co-precipitated with GFP::GLR-1 (Fig. 5d).

Finally, we examined the topological arrangement of SOL-1. Sequence analysis predicted a large extracellular N-terminal domain. To test this, we dissociated neurons from transgenic worms expressing SOL-1 tagged with GFP at the N terminus. Under non-permeabilized conditions, antibodies directed against GFP detected cell-surface expression of GFP::SOL-1, suggesting that the N terminus is extracellular (Fig. 5e–h). To confirm this result, we injected fluorescently labelled anti-GFP antibodies into transgenic worms expressing either N-terminal or C-terminal GFP-tagged SOL-1. The antibody label was observed only in worms expressing the N-terminal tagged SOL-1 (Fig. 5i, j). Thus, our results indicate that SOL-1 is a type I transmembrane protein that is expressed at the cell surface.

Most GLR-1 receptor oligomers also contain GLR-2 subunits¹². Mutations in *glr-2* disrupt nose touch avoidance, but to a lesser extent than do mutations in *glr-1* (ref. 12) or *sol-1*. Antibody labelling of GFP-tagged GLR-2 was unaffected in *sol-1* mutants, indicating that GLR-2 trafficking and surface expression does not depend on SOL-1 (Supplementary Fig. S3). Furthermore, it has been shown that a fast, yet small, GLR-1-dependent glutamate-gated current is present even in the absence of GLR-2 (ref. 12). Because we did not observe this small, GLR-1-dependent, fast glutamate-gated current in *sol-1* mutants (Fig. 4b), we conclude that SOL-1 is not specifically required for the expression or

assembly of GLR-2-containing heteromeric receptors.

Using a genetic strategy we have identified SOL-1, a gene product that is required for GLR-1-dependent glutamate-gated currents and for behaviours that are dependent on GLR-1-mediated synaptic transmission^{7,8,12}. The evidence suggests that SOL-1 and GLR-1 are part of a receptor complex: SOL-1 colocalizes with GLR-1 at synaptic puncta, and these proteins physically interact when expressed in heterologous cells. SOL-1 may directly influence GLR-1 function or it may recruit other proteins that interact with GLR-1. Sequence analysis identified four protein–protein interaction CUB domains but no other predicted domains. CUB domain proteins have been identified in the brain^{19,20}; these same molecules and others, such as neuropilin^{21,22}, share sequence identity with SOL-1 in the CUB domain regions²³.

We suggest that SOL-1 is required for gating of the receptor because the membrane expression of GLR-1 receptors is not altered in *sol-1* mutants, and because it is doubtful that ligand binding is modified as the soluble S1–S2 extracellular domain of GluR2 binds glutamate and kainate with near normal affinity²⁴. SOL-1 may ensure that GLR-1 iGluRs are not functional if, for example, they are located in intracellular organelles or are targeted to inappropriate regions of the cell surface. SOL-1 may also regulate the gain of the postsynaptic response by controlling the number of functional receptors—perhaps acting as a so-called ‘slot’ protein similar to those proposed to regulate the density of functional synaptic iGluRs²⁵. We anticipate that homologues of SOL-1 may regulate

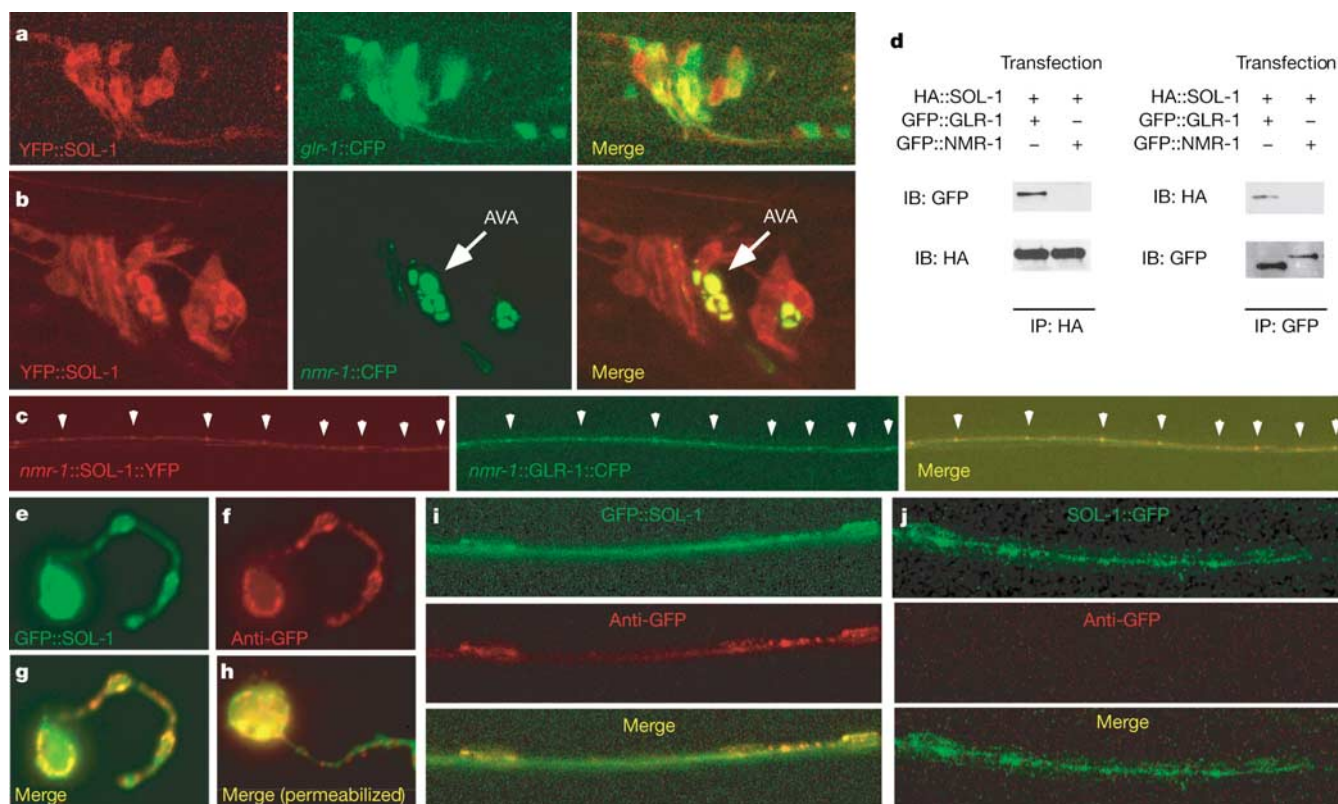


Figure 5 SOL-1 is a predicted type I transmembrane protein that interacts with GLR-1. **a, b**, Confocal images of the head region of transgenic worms expressing YFP::SOL-1 (left) and CFP (middle) driven by either the *glr-1* (**a**; *lin-15(n765ts)*; *akEx246*) or the *nmr-1* (**b**; *lin-15(n765ts)*; *akEx247*) promoter. Merged image (right). **c**, Confocal images of ventral cord processes in a transgenic worm expressing SOL-1::YFP (left) and GLR-1::CFP (middle) under control of the *nmr-1* promoter. Arrowheads indicate YFP and CFP puncta, which completely overlap in the merged image (right). **d**, Co-immunoprecipitation of HA::SOL-1 and GFP::GLR-1 from COS-7 cells. Cells were transfected with HA::SOL-1 and

either GFP::GLR-1 or GFP::NMR-1. **e-h**, Neurons expressing GFP::SOL-1 cultured from transgenic worms. Shown are GFP::SOL-1 (**e**), anti-GFP staining (**f**) and the merged image (**g**) in a non-permeabilized neuron, and the merged image of GFP::SOL-1 and anti-GFP staining in a permeabilized neuron (**h**). **i, j**, Confocal images of GFP (top), anti-GFP staining (middle) and the merged images (bottom) in transgenic worms expressing GFP::SOL-1 (N-terminal fusion, **i**) or SOL-1::GFP (C-terminal fusion, **j**) injected with Alexa594-conjugated rabbit anti-GFP antibodies.

iGluRs in the vertebrate nervous system and contribute to synaptic plasticity. □

Methods

Strains, genetics and germline transformation

All strains were raised at 20 °C under standard conditions. Germline transformation was achieved as described¹⁶ using the *lin-15* clone pJM23 (40 ng μ l⁻¹) as a transformation marker. An F₂ screen was used to isolate recessive mutations that suppressed the hyper-reversal behaviour of *lurcher* (*akIs9*) worms⁸. *akIs9* worms were mutagenized with 50 mM EMS at the L4 larval stage using standard techniques²⁶. The F₂ progeny (representing 12,000 haploid genomes) of the mutagenized worms were washed off NGM plates in M9 buffer, placed in a small drop of M9 at one pole of a 10-cm NGM plate and allowed to migrate towards a food source (*Escherichia coli* strain OP50) at the opposite pole. Worms that reached the food within 1 h were isolated as potential suppressors of the *akIs9* hyper-reversal phenotype.

We cloned *sol-1* by standard genetic mapping techniques, followed by transformation rescue of *sol-1(ak63)*; *akIs9* worms. We outcrossed *sol-1* alleles at least three times (*ak63*, five times). *sol-1(ak63)*; *lin-15(n765ts)* or *lin-15(n765ts)* mutants were used in all transgenic experiments and expressed the following extrachromosomal arrays: *akEx204*, pYZ96; *akEx241*, pYZ134; *akEx243*, pYZ128; *akEx246*, pYZ138 + pYZ112; and *akEx247*, pYZ138 + pYZ113 (all arrays contained pJM23 to rescue the *lin-15* phenotype⁷). The nature of the three *sol-1* mutant alleles (*ak63*, *ak66* and *ak67*) was determined by sequencing multiple PCR products amplified from genomic DNA isolated from each of the mutant strains. We isolated the *sol-1* cDNA by PCR amplification from the *C. elegans* first strand cDNA (GenBank accession number AY387697). This cDNA differed slightly from that predicted by the Genefinder analysis program²⁷. The predicted SOL-1 protein was analysed by the ExPASy suite of programs²⁸.

Plasmid constructs

The *sol-1* genomic clone was constructed by subcloning a 12.3-kb *XmaI*-*StuI* fragment

from C15A11 into pBluescript SK+ (Stratagene). N-terminal GFP::SOL-1 and C-terminal SOL-1::GFP were generated by inserting GFP after the third amino acid after the predicted signal sequence, and immediately upstream of the predicted *sol-1* stop codon, respectively. For expression in COS-7 cells, GFP and HA coding sequences were fused in frame after the third amino acid following the predicted signal sequence for the *glr-1* and *nmr-1* cDNAs, and for the *sol-1* cDNA, respectively, and then subcloned into the mammalian expression vector pCDNA3.1(+) (Invitrogen). The plasmids were as follows: pYZ96, *sol-1* genomic clone; pYZ134, GFP::SOL-1; pYZ152, *nmr-1*::GFP::SOL-1; pYZ138, YFP::SOL-1; pYZ112, *glr-1*::CFP; pYZ113, *nmr-1*::CFP; pYZ167, *nmr-1*::YFP::SOL-1; pYZ165, *nmr-1*::GLR-1::CFP; pYZ128, SOL-1::GFP; pYZ261, HA::SOL-1; pDM439, GFP::GLR-1; pDM472, GFP::NMR-1.

Primary cultures of *C. elegans* neurons

Neurons were cultured from transgenic strains that expressed GFP fused to full-length GLR-1 or SOL-1 by using described techniques¹⁵. We used the following strains: KP1477, *GLR-1::GFP* (*nuls25*)¹⁴; VM1557, *nuls25 sol-1(ak63)*; VM1447, *sol-1(ak63)*; *akEx241*; VM1472, *sol-1(ak63)*; *akEx243*.

Immunolabelling and microscopy

Immunolabelling was done in both cultured neurons and in live worms. Primary cultured neurons were fixed for 30 min at room temperature with 3% formaldehyde. For analysis of GLR-1 or GFP::SOL-1 cell-surface expression, cells were labelled with either affinity-purified rabbit anti-GLR-1 (ref. 9) or goat anti-GFP polyclonal sera followed by Cy3-conjugated anti-rabbit or anti-goat secondary antibodies in the presence (permeabilized conditions) or absence (non-permeabilized conditions) of 0.2% Triton X-100. Epifluorescence images were acquired with an Axiocvert 135 microscope (Zeiss) and a Micromax CCD camera (Princeton Instruments).

Immunolabelling in live worms was done using a previously developed method (A. Gottschalk and W. Schafer, personal communication). In brief, Alexa594-conjugated rabbit anti-GFP polyclonal sera (Molecular Probes) were diluted (1:200) in worm injection buffer and injected into the pseudocoelom of transgenic worms expressing one of the following four chimaeras: GFP::GLR-1, GLR-1::GFP, GFP::SOL-1 or

SOL-1::GFP. To clear excess antibody from the pseudocoelomic space, we allowed injected worms to recover for 6 h on NGM plates seeded with the *E. coli* strain OP50. All labelling experiments were repeated at least four times. Epifluorescence and GFP images were acquired using a LSM 510 confocal imaging system (Zeiss).

Tissue culture and immunoprecipitation

COS-7 cells were cultured in DMEM nutrient medium with 10% bovine fetal serum. The cells were transiently transfected by using Lipofectamine2000 (Invitrogen), collected after 36 h, and lysed in ice-cold IP buffer (25 mM Tris (pH 7.5), 150 mM NaCl, 10% glycerol, 1% Triton X-100 and protease inhibitor cocktail (Roche)). The lysates were incubated with agarose-conjugated antibodies (Santa Cruz Biotech) in IP buffer overnight at 4°C. After three washes in ice-cold IP buffer, the samples were boiled in SDS-PAGE sample buffer for 2 min, and the precipitated proteins were resolved on a 7% SDS-PAGE gel, probed with peroxidase-conjugated antibodies (Jackson ImmunoResearch) and analysed with the Supersignal Chemiluminescence kit (Pierce).

Behavioural and electrophysiological studies

Drug delivery was done by pressure application using a Picospritzer II (General Valve Corporation). We made electrophysiological recordings of ligand-gated currents in the AVA interneurons using standard patch-clamp technology as described¹⁶. All electrophysiological experiments were repeated 3–8 times. Behavioural assays for the nose touch response, osmotic avoidance and the mean duration of forward and backward movement have been described¹². Statistical significance was determined using the standard Student's *t*-test. Error bars represent the s.e.m.

Received 23 September; accepted 20 November 2003; doi:10.1038/nature02244.

1. Zuo, J. *et al.* Neurodegeneration in Lurcher mice caused by mutation in $\delta 2$ glutamate receptor gene. *Nature* **388**, 769–773 (1997).
2. Bork, P. & Beckmann, G. The CUB domain. A widespread module in developmentally regulated proteins. *J. Mol. Biol.* **231**, 539–545 (1993).
3. Varela, P. F. *et al.* The 2.4 Å resolution crystal structure of boar seminal plasma PSP-I/PSP-II: a zona pellucida-binding glycoprotein heterodimer of the spermadhesin family built by a CUB domain architecture. *J. Mol. Biol.* **274**, 635–649 (1997).
4. Dingleline, R., Borges, K., Bowie, D. & Traynelis, S. F. The glutamate receptor ion channels. *Pharmacol. Rev.* **51**, 7–61 (1999).
5. Sheng, M. & Kim, M. J. Postsynaptic signaling and plasticity mechanisms. *Science* **298**, 776–780 (2002).
6. Kohda, K., Wang, Y. & Yuzaki, M. Mutation of a glutamate receptor motif reveals its role in gating and $\delta 2$ receptor channel properties. *Nature Neurosci.* **3**, 315–322 (2000).
7. Maricq, A. V., Peckol, E., Driscoll, M. & Bargmann, C. I. Mechanosensory signalling in *C. elegans* mediated by the GLR-1 glutamate receptor. *Nature* **378**, 78–81 (1995).
8. Hart, A. C., Sims, S. & Kaplan, J. M. Synaptic code for sensory modalities revealed by *C. elegans* GLR-1 glutamate receptor. *Nature* **378**, 82–85 (1995).
9. Zheng, Y., Brockie, P. J., Mellem, J. E., Madsen, D. M. & Maricq, A. V. Neuronal control of locomotion in *C. elegans* is modified by a dominant mutation in the GLR-1 ionotropic glutamate receptor. *Neuron* **24**, 347–361 (1999).
10. Stein, L., Sternberg, P., Durbin, R., Thierry-Mieg, J. & Spieth, J. WormBase: network access to the genome and biology of *Caenorhabditis elegans*. *Nucleic Acids Res.* **29**, 82–86 (2001).
11. Gregory, L. A., Thielens, N. M., Arlaud, G. J. & Fontecilla-Camps, J. C. Gaboriaud C. X-ray structure of the Ca^{2+} -binding interaction domain of C1s: insights into the assembly of the C1 complex of complement. *J. Biol. Chem.* **278**, 32157–32164 (2003).
12. Mellem, J. E., Brockie, P. J., Zheng, Y., Madsen, D. M. & Maricq, A. V. Decoding of polymodal sensory stimuli by postsynaptic glutamate receptors in *C. elegans*. *Neuron* **36**, 933–944 (2002).
13. Hilliard, M. A., Bargmann, C. I. & Bazzicalupo, P. C. *C. elegans* responds to chemical repellents by integrating sensory inputs from the head and the tail. *Curr. Biol.* **12**, 730–734 (2002).
14. Rongo, C., Whitfield, C. W., Rodal, A., Kim, S. K. & Kaplan, J. M. LIN-10 is a shared component of the polarized protein localization pathways in neurons and epithelia. *Cell* **94**, 751–759 (1998).
15. Christensen, M. *et al.* A primary culture system for functional analysis of *C. elegans* neurons and muscle cells. *Neuron* **33**, 503–514 (2002).
16. Brockie, P. J., Mellem, J. E., Hills, T., Madsen, D. M. & Maricq, A. V. The *C. elegans* glutamate receptor subunit NMR-1 is required for slow NMDA-activated currents that regulate reversal frequency during locomotion. *Neuron* **31**, 617–630 (2001).
17. Brockie, P. J., Madsen, D. M., Zheng, Y., Mellem, J. & Maricq, A. V. Differential expression of glutamate receptor subunits in the nervous system of *Caenorhabditis elegans* and their regulation by the homeodomain protein UNC-42. *J. Neurosci.* **21**, 1510–1522 (2001).
18. Chelur, D. S. *et al.* The mechanosensory protein MEC-6 is a subunit of the *C. elegans* touch-cell degenerin channel. *Nature* **420**, 669–673 (2002).
19. Michishita, M. *et al.* A novel gene, Btcl1, encoding CUB and LDLA domains is expressed in restricted areas of mouse brain. *Biochem. Biophys. Res. Commun.* **306**, 680–686 (2003).
20. Stohr, H., Berger, C., Fröhlich, S. & Weber, B. H. A novel gene encoding a putative transmembrane protein with two extracellular CUB domains and a low-density lipoprotein class A module: isolation of alternatively spliced isoforms in retina and brain. *Gene* **286**, 223–231 (2002).
21. He, Z. & Tessier-Lavigne, M. Neuropilin is a receptor for the axonal chemorepellent Semaphorin III. *Cell* **90**, 739–751 (1997).
22. Kolodkin, A. L. *et al.* Neuropilin is a semaphorin III receptor. *Cell* **90**, 753–762 (1997).
23. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
24. Armstrong, N., Sun, Y., Chen, G. Q. & Gouaux, E. Structure of a glutamate-receptor ligand-binding core in complex with kainate. *Nature* **395**, 913–917 (1998).
25. Shi, S., Hayashi, Y., Esteban, J. A. & Malinow, R. Subunit-specific rules governing AMPA receptor trafficking to synapses in hippocampal pyramidal neurons. *Cell* **105**, 331–343 (2001).
26. Anderson, P. in *Methods in Cell Biology* (eds Epstein, H. F. & Shakes, D. C.) 31–58 (Academic, New York, 1995).

27. Wilson, R. *et al.* 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. *Nature* **368**, 32–38 (1994).
28. Gasteiger, E. *et al.* ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784–3788 (2003).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank M. Vetter and members of the Maricq laboratory for comments on the manuscript; L. Jack for generating transgenic strains; C. Walker, N. Strutz, M. Francis and A. Ebens for discussions; C. Rongo and J. Kaplan for the *nuls25* strain; and A. Gottschalk and W. Schafer for help with immunolabelling live worms. Some strains were provided by the *Caenorhabditis* Genetics Center. This research was supported by the Burroughs Wellcome Foundation, and by a grant from the NIH.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to A.V.M. (maricq@biology.utah.edu).

The RickA protein of *Rickettsia conorii* activates the Arp2/3 complex

Edith Gouin¹, Coumaran Egile², Pierre Dehoux¹, Véronique Villiers¹, Josephine Adams³, Frank Gertler⁴, Rong Li² & Pascale Cossart¹

¹Unité des Interactions Bactéries-Cellules, Institut Pasteur, 28 Rue du Docteur Roux, Paris 75015, France

²Harvard Medical School, Department of Cell Biology, Boston, Massachusetts 02115, USA

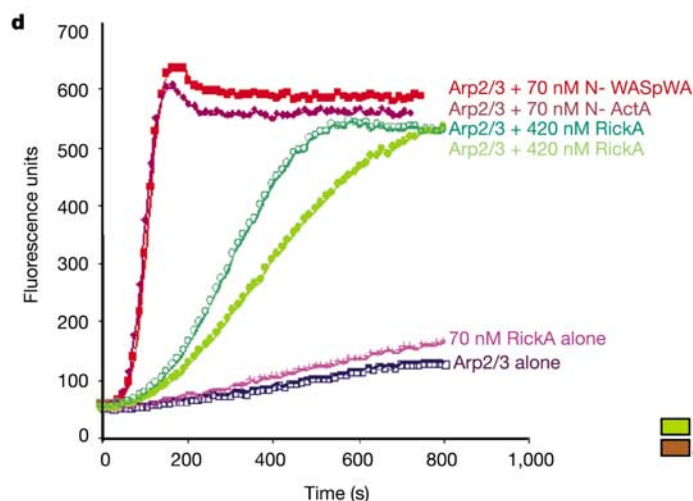
³Cleveland Clinic Foundation, Department of Cell Biology, Cleveland, Ohio 44195, USA

⁴Massachusetts Institute of Technology, Department of Biology, Cambridge, Massachusetts 02139, USA

Actin polymerization, the main driving force for cell locomotion, is also used by the bacteria *Listeria* and *Shigella* and vaccinia virus for intracellular and intercellular movements^{1,2}. Seminal studies have shown the key function of the Arp2/3 complex in nucleating actin and generating a branched array of actin filaments during membrane extension and pathogen movement³. Arp2/3 requires activation by proteins such as the WASP-family proteins or ActA of *Listeria*. We previously reported that actin tails of *Rickettsia conorii*, another intracellular bacterium, unlike those of *Listeria*, *Shigella* or vaccinia, are made of long unbranched actin filaments apparently devoid of Arp2/3 (ref. 4). Here we identify a *R. conorii* surface protein, RickA, that activates Arp2/3 *in vitro*, although less efficiently than ActA. In infected cells, Arp2/3 is detected on the rickettsial surface but not in actin tails. When expressed in mammalian cells and targeted to the membrane, RickA induces filopodia. Thus RickA-induced actin polymerization, by generating long actin filaments reminiscent of those present in filopodia, has potential as a tool for studying filopodia formation.

Converging studies on actin-based motility have led to the establishment of the essential role of Arp2/3 in this process^{1–3}. Arp2/3 activates the nucleation of actin polymerization and induces the continuous formation of a network of short and highly branched filaments in lamellipodia and pathogen actin tails^{1–3}. Arp2/3 is normally inactive and has to be activated by proteins of the WASP/N-WASP/Scar/Wave-family proteins in eukaryotic cells or the ActA protein at the surface of *L. monocytogenes*^{1,3–8}. The *Shigella* protein VirG/IcsA and the vaccinia viral protein A36R are not directly involved in the activation of nucleation². They recruit N-WASP and then Arp2/3, respectively directly and indirectly. Other known activators of the Arp2/3 complex include yeast

Members of the genus *Rickettsia* are obligate intracellular bacteria that grow in the cytoplasm of their eukaryotic host¹³. They are responsible for several serious human diseases including epidemic typhus (*Rickettsia prowazekii*) and spotted fevers (*Rickettsia conorii*,



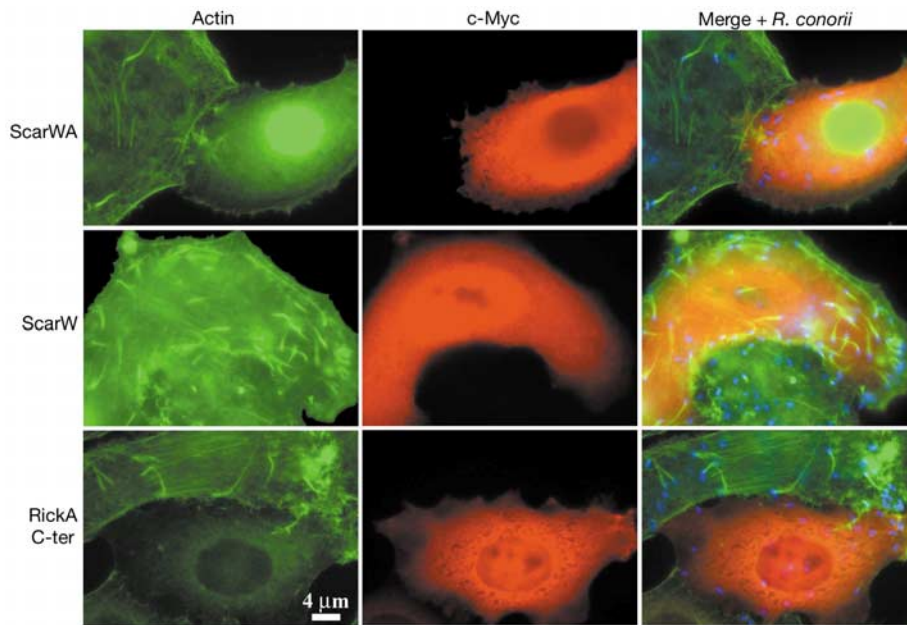


Figure 2 Transfection with ScarWA or the C-terminal part of RickA inhibits actin polymerization. Hep-2 cells were infected with *R. conorii* 24 h before transfection with the plasmids encoding ScarWA, ScarW or RickA-C-ter (the fragment containing residues 377–517 from RickA was cloned in pRK5Myc (ref. 28), resulting in plasmid pRC19 (strain

BUG1998)). The cells were processed for immunofluorescence 24 h after transfection with fluorescein isothiocyanate-conjugated phalloidin for F-actin, an anti-c-Myc monoclonal antibody for transfected cells and the polyclonal R47 antibody⁴ for *R. conorii*.

known to allow binding of the Ena/VASP proteins, surrounded by an amino-terminal domain with a potential G-actin-binding site, similar to that of thymosin β 4 (ref. 17), and a carboxy-terminal region that shares similarities with the CA domains of WASP-family proteins (Fig. 1a and Supplementary Fig. 1), including the conserved hydrophobic domain and the positively charged region that would form the predicted amphipathic helix able to bind the Arp2/3 complex¹⁸. The amino-acid sequence of RickA does not display any signal sequence or a C-terminal motif that could act as a membrane anchor. Interestingly, the RickA paralogues in *R. montana*, *R. rickettsii* and *R. siberica* have the same general organization as RickA but with some variation in the number and the amino acid sequences of the proline-rich repeats (data not shown).

To study the function of the protein, we produced, although with very low yields, a recombinant protein in *Escherichia coli* and also raised antibodies against a peptide present in the N-terminal region of RickA. As shown by immunofluorescence (Fig. 1b), the RickA protein is highly expressed on the surface of all *Rickettsia* spp., in infected cells. In addition, western blot analysis showed that the RickA protein, present in *Rickettsia* extracts and absent from extracts of non-infected cells, migrated as both a protein of 70 kDa (as does the recombinant protein) and a protein of 140 kDa, indicating that RickA can dimerize (Fig. 1c).

Actin polymerization assays *in vitro* with the N-terminal part of ActA (N-ActA) as a control showed that RickA alone was unable to stimulate actin polymerization but was able to activate the Arp2/3 complex, in a dose-dependent manner and as efficiently as ActA (Fig. 1d). However, in contrast to ActA, RickA did not reduce the lag period preceding the exponential phase of actin nucleation.

To investigate further the unexpected role of RickA in Arp2/3 activation we also used the technique¹⁹ that analyses the capacity of activated Arp2/3 to induce branch formation in actin filaments. Again taking N-ActA as a control, we showed that in the presence of actin and Arp2/3 the RickA protein promotes filament branching (Fig. 1e). However, a similar percentage of branching, at equivalent concentrations of Arp2/3 and actin, required sixfold more RickA

than ActA and was reached after 15 min for RickA but after only 1 min for ActA (Fig. 1e).

Having shown that RickA is sufficient to activate Arp2/3, and thereby actin nucleation *in vitro*, it was of interest to clarify whether the Arp2/3 complex has a function in the formation of the *Rickettsia* actin tail and in the actin-based motility process *in vivo*, and also whether it is present in the tails.

We therefore first transfected cells with plasmids expressing either the C-terminal part of Scar (WA)—which can recruit and sequester Arp2/3 and acts as a dominant-negative fragment with respect to the function of the Arp2/3 complex—or a truncated version (W), and infected these cells with *Rickettsia* (Fig. 2), using *Listeria* as a control (data not shown)⁸. Clearly, *Rickettsia* and *Listeria* actin tails were both detected in ScarW-transfected cells. In contrast, the number of bacteria (*Rickettsia* or *Listeria*) recruiting actin was greatly decreased in ScarWA-expressing cells, further indicating that the Arp2/3 complex might be used by *Rickettsia*, as it is in *Listeria*, to induce actin polymerization and movement.

We next re-examined Arp2/3 recruitment in the tails. All previous attempts to detect the Arp2/3 complex in *Rickettsia* tails had failed^{4,20}. Using a new batch of affinity-purified antibodies, we were able to detect Arp3 around *Rickettsia*. Strikingly, Arp3 localized with cytoplasmic bacteria, whether with or without actin tails (Fig. 3). However, for those without actin tails (arrowheads in Fig. 3, top panels), Arp3 was detected around the whole bacterial bodies, and the strong labelling indicates that it was probably abundantly recruited. In contrast, when bacteria were polymerizing actin (arrows in Fig. 3, upper panels), Arp3 was barely detectable at the base of the tail and was absent from the tail. This distribution of Arp3 is therefore different from that in *Listeria* tails (arrowheads in Fig. 3, lower panels), in which Arp3 localizes with actin and is present both at the base and all along the tail. These observations are in complete agreement with the *in vitro* data and strongly indicate that RickA can recruit and activate the Arp2/3 complex to induce actin polymerization in infected cells.

If RickA recruits the Arp2/3 complex, it could behave as

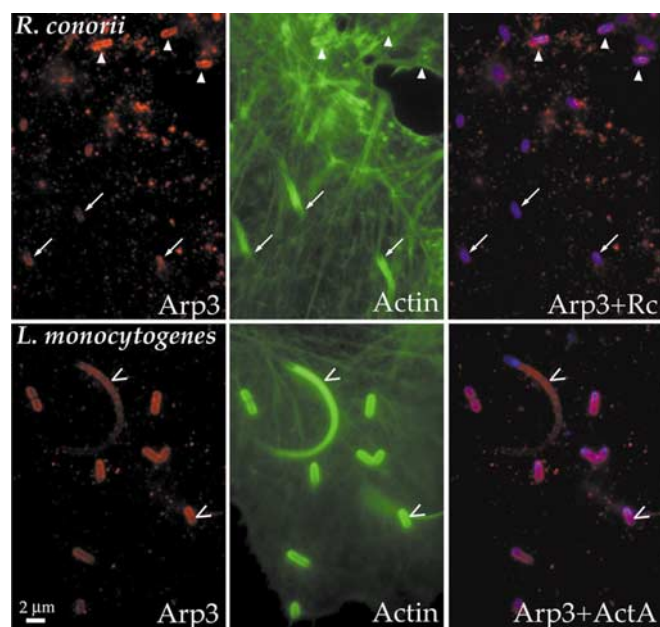


Figure 3 Arp3 localizes with the bacteria *R. conorii* and with *L. monocytogenes* comet tails in Hep-2-infected cells. Upper panels, Hep-2 cells infected with *R. conorii*; lower panels, Hep-2 cells infected with *L. monocytogenes*. A new batch of affinity-purified polyclonal anti-Arp3 (0.94 mg ml⁻¹), described in ref. 29, was used. *L. monocytogenes* cells were labelled with monoclonal anti-N-ActA A16 (ref. 6).

ScarWA when overexpressed in cells. To test this hypothesis we expressed the C-terminal part of RickA, which is similar to the VCA domain of N-WASP, in cells and infected them with *Listeria* or *Rickettsia* (Fig. 2). The results were similar to those obtained with ScarWA: transfection with the C-terminal part of RickA impaired *Rickettsia* tail formation. For *Listeria*, bacterial entry, which is known to be dependent on Arp2/3, was even more strongly inhibited than with ScarWA. Collectively, both the *in vitro* and *in vivo* data demonstrate that RickA is a previously unknown Arp2/3 activator.

Because the known Arp2/3 activators have been involved in cell motility or the formation of plasma membrane protrusions, we were interested to analyse how RickA would behave when expressed at the plasma membrane. We therefore transfected cells with a plasmid expressing a RickA-CAAX construct designed to drive expression of the protein at the inner face of the plasma membrane, as previously performed for ActA²¹. As shown in Fig. 4, after RickA expression, Vero-transfected cells displayed thin filopodia that were absent from non-transfected cells. These filopodia, whose length varied with the cell line used for the transfection, were different from the lamellipodia-like structures obtained when ActA, or even more strikingly IcsA, was transfected in mammalian cells, reinforcing the hypothesis that RickA is a novel type of activator of actin nucleation.

In conclusion, *Rickettsia conorii* cells express a surface protein, RickA, which recruits Arp2/3, activates it and induces actin polymerization. It is unknown how RickA is addressed to the bacterial surface and whether the type IV secretion system predicted by the genome sequence is involved in targeting to the surface. The actin filaments present behind intracellular bacteria in the tails are long and unbranched (Supplementary Fig. 2). These filaments might form bundles, in particular when bacteria spread from cell to cell, as shown previously⁴. Accordingly, as filopodia, *Rickettsia* tails contain the bundling protein fascin (Supplementary Fig. 3)²². How could such bundles form? In agreement with the well-established findings that activated Arp2/3 initiates actin polymerization on previously

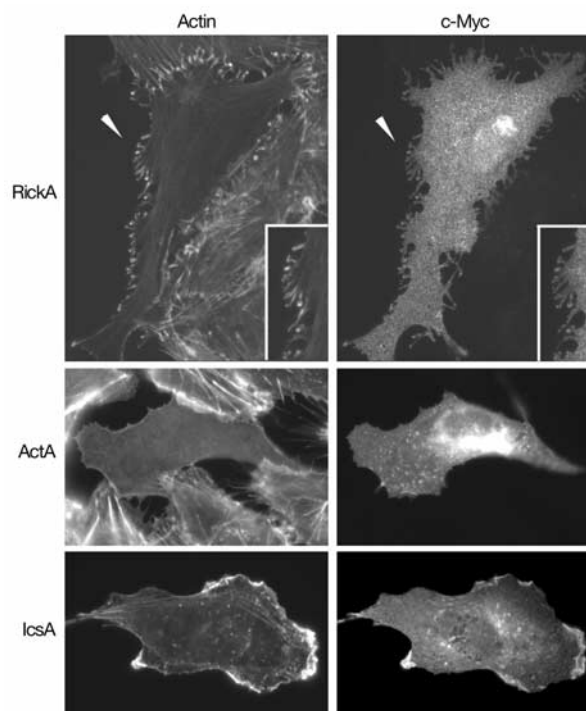


Figure 4 Transfection of mammalian cells with RickA-CAAX induces filopodia-like structure. Vero cells were transfected with plasmids expressing RickA-CAAX, ActA-CAAX or IcsA-CAAX, and processed for immunofluorescence 24 h later. Arrowheads point to regions shown in the inserts.

formed actin filaments, generating a branch, it is possible that RickA initiates actin polymerization in such a way that actin filaments nucleated by Arp2/3 are then elongated at only one barbed end, just as in filopodia¹². Could VASP have a function in this process? VASP is present all along the *Rickettsia* tail⁴ but is restricted to the base of the *Listeria* tails, where it binds directly to ActA through its EVH1 domain²³. We infected D7 cells, which are completely devoid of Mena, VASP and EVL²⁴, with *Rickettsia* or *Listeria* and compared the bacterial motility in these cells with that observed in D7 cells transfected with Mena. *Rickettsia* cells were less motile in the D7 cells, as were *Listeria* cells, suggesting that VASP-family proteins are required for maximal movement in *Rickettsia*. In both lamellipodia²⁵ and ActA-induced actin assembly²⁶, VASP decreases branch density and, as proposed, could compete with capping proteins at the barbed ends²⁵ and thus increase filament length. This could also occur in RickA-induced actin assembly. In *Shigella*, Ena/VASP proteins are located throughout the tails, but reconstitution experiments suggest that Ena/Vasp proteins are dispensable for *Shigella* motility⁷. The role of VASP in *Rickettsia* motility therefore seems different from that in *Listeria* and *Shigella* motilities.

Taken together, the *Rickettsia* protein RickA is a previously unknown bacterial activator of actin nucleation that requires Arp2/3 to initiate an actin polymerization process similar to that leading to filopodia¹². Whether proteins thought to contribute to filopodia formation and/or predicted to be present at the tip of the filopodia are also involved in *Rickettsia* movement is unknown. It is possible that post-translational modifications of RickA such as phosphorylations at various positions including Ser 459, a position corresponding to Ser 484 in WASP²⁷, controls its activity. Whether another *Rickettsia* protein also participates in the actin-based motility has to be addressed. Thus, RickA is a bacterial actin nucleator that is most closely related to WASP/N-WASP-family

proteins. It is unknown whether RickA is of eukaryotic origin or has been transferred from *Rickettsia* to eukaryotic cells. □

Methods

Expression and purification of recombinant proteins

RickA, the ActA-N fragment (residues 1–233) and the WA fragment of N-WASP (residues 392–505) were obtained as His-tagged proteins. The *rickA* gene was amplified by polymerase chain reaction with *Rickettsia conorii* (Malish strain) chromosomal DNA as template and the primers 5'-CCatggttaagaagaatagataataaaa-3' and 5'-caagctttctaacaatgatgggtttt-3'. This fragment was digested with *NcoI* and *HindIII* and cloned into these sites in pET28b. The DNA fragment encoding the WA fragment of N-WASP was cloned into the *EcoRI* and *XhoI* sites of pET28a to generate an N-terminal His₆-tagged protein. Proteins were expressed in *E. coli* BL21(DE3) and affinity purified with Talon resin (Clontech).

Actin polymerization assays

Polymerization assays were performed with the change in pyrenyl-actin fluorescence. Pyrene-labelled rabbit muscle actin (10% pyrene-labelled) in G buffer (0.2 mM CaCl₂, 0.2 mM DTT, 0.2 mM ATP, 1 mM MgCl₂, 5 mM Tris-HCl pH 7.5) was precleared at 250,000 g for 30 min. MgATP-actin (1.5 μM, 10% pyrene-labelled) was polymerized in 1 × KMET (50 mM KCl, 1 mM MgCl₂, 1 mM EGTA, 10 mM Tris-HCl pH 7) supplemented with 0.5 mM ATP and RickA (70 nM) or Arp2/3 (30 nM) alone or in the presence of N-WASP-WA (70 nM), N-ActA (70 nM) and RickA (70 and 420 nM). Measurements were made with a Varian Eclipse Spectrofluorometer at 25 °C, with excitation and emission wavelengths of 350 and 390 nm respectively.

Membrane targeting of RickA-CAAX, ActA-CAAX and IcsA-CAAX

To express RickA at the plasma membrane, an approach similar to that used in ref. 21 was taken with plasmid pRK5Myc (ref. 28), which allows the expression of N-terminal Myc-tagged and C-terminal CAAX-tagged protein. DNA fragments encoding ActA (1–580), RickA (1–517) and IcsA (53–758) were cloned in the *BamHI* and *EcoRI* sites of pRK5Myc-CAAX. The resultant plasmids were transiently expressed in Vero cells with Lipofectamine 2000.

Antibodies

Anti-RickA antibodies (R80) were generated by immunizing rabbits against the peptide CQNETKELEKEHNRS and affinity purified. Anti-*Rickettsia* mouse polyclonal antibodies (S1) were obtained by immunizing BALBc mice with formalin-killed bacteria purified from infected yolk sac.

Received 28 October; accepted 23 December 2003; doi:10.1038/nature02318.

- Pollard, T. D. & Borisy, G. G. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* **112**, 453–465 (2003).
- Frischknecht, F. & Way, M. Surfing pathogens and the lessons learned for actin polymerization. *Trends Cell Biol.* **11**, 30–38 (2001).
- Welch, M. D., Rosenblatt, J., Skoble, J., Portnoy, D. A. & Mitchison, T. J. Interaction of human Arp2/3 complex and the *Listeria monocytogenes* ActA protein in actin filament nucleation. *Science* **281**, 105–108 (1998).
- Gouin, E. *et al.* A comparative study of the actin-based motilities of the pathogenic bacteria *Listeria monocytogenes*, *Shigella flexneri* and *Rickettsia conorii*. *J. Cell Sci.* **112**, 1697–1708 (1999).
- Machesky, L. M. *et al.* Scar, a WASP-related protein, activates nucleation of actin filaments by the Arp2/3 complex. *Proc. Natl Acad. Sci. USA* **96**, 3739–3744 (1999).
- Boujemaa-Paterski, R. *et al.* *Listeria* protein ActA mimics WASP family proteins: it activates filament barbed end branching by Arp2/3 complex. *Biochemistry* **40**, 11390–11404 (2001).
- Loisel, T. P., Boujemaa, R., Pantaloni, D. & Carlier, M. F. Reconstitution of actin-based motility of *Listeria* and *Shigella* using pure proteins. *Nature* **401**, 613–616 (1999).
- May, R. C. *et al.* The Arp2/3 complex is essential for the actin-based motility of *Listeria monocytogenes*. *Curr. Biol.* **9**, 759–762 (1999).
- Winter, D., Lechler, T. & Li, R. Activation of the yeast Arp2/3 complex by Bee1p, a WASP-family protein. *Curr. Biol.* **9**, 501–504 (1999).
- Weaver, A. M. *et al.* Cortactin promotes and stabilizes Arp2/3-induced actin filament network formation. *Curr. Biol.* **11**, 370–374 (2001).
- Uruno, T. *et al.* Activation of Arp2/3 complex-mediated actin polymerization by cortactin. *Nature Cell Biol.* **3**, 259–266 (2001).
- Vignjevic, D. *et al.* Formation of filopodia-like bundles *in vitro* from a dendritic network. *J. Cell Biol.* **160**, 951–962 (2003).
- Hackstadt, T. The biology of *Rickettsia*. *Infect. Agents Dis.* **5**, 127–143 (1996).
- Teyssie, N., Chiche-Portiche, C. & Raoult, D. Intracellular movements of *Rickettsia conorii* and *R. typhi* based on actin polymerization. *Res. Microbiol.* **143**, 821–829 (1992).
- Andersson, S. G. *et al.* The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria. *Nature* **396**, 133–140 (1998).
- Ogata, H. *et al.* Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. *Science* **293**, 2093–2098 (2001).
- Van Troys, M. *et al.* The actin binding site of thymosin beta 4 mapped by mutational analysis. *EMBO J.* **15**, 201–210 (1996).
- Panchal, S. C., Kaiser, D. A., Torres, E., Pollard, T. D. & Rosen, M. K. A conserved amphipathic helix in WASP/Scar proteins is essential for activation of Arp2/3 complex. *Nature Struct. Biol.* **10**, 591–598 (2003).
- Blanchoin, L. *et al.* Direct observation of dendritic actin filament networks nucleated by Arp2/3 complex and WASP/Scar proteins. *Nature* **404**, 1007–1011 (2000).
- Van Kirk, L. S., Hayes, S. F. & Heinzen, R. A. Ultrastructure of *Rickettsia rickettsii* actin tails and localization of cytoskeletal proteins. *Infect. Immun.* **68**, 4706–4713 (2000).

- Friedrich, E. *et al.* Targeting of *Listeria monocytogenes* ActA protein to the plasma membrane as a tool to dissect both actin-based cell morphogenesis and ActA function. *EMBO J.* **14**, 2731–2744 (1995).
- Kureishy, N., Sapountzi, V., Prag, S., Anilkumar, N. & Adams, J. C. Fascins, and their roles in cell structure and function. *Bioessays* **24**, 350–361 (2002).
- Chakraborty, T. *et al.* A focal adhesion factor directly linking intracellularly motile *Listeria monocytogenes* and *Listeria ivanovii* to the actin-based cytoskeleton of mammalian cells. *EMBO J.* **14**, 1314–1321 (1995).
- Bear, J. E. *et al.* Negative regulation of fibroblast motility by Ena/VASP proteins. *Cell* **101**, 717–728 (2000).
- Bear, J. E. *et al.* Antagonism between Ena/VASP proteins and actin filament capping regulates fibroblast motility. *Cell* **109**, 509–521 (2002).
- Skoble, J. *et al.* Pivotal role of VASP in Arp2/3 complex-mediated actin nucleation, actin branch formation, and *Listeria monocytogenes* motility. *J. Cell Biol.* **155**, 89–100 (2001).
- Cory, G. O., Cramer, R., Blanchoin, L. & Ridley, A. J. Phosphorylation of the WASP-VCA domain increases its affinity for the Arp2/3 complex and enhances actin polymerization by WASP. *Mol. Cell* **11**, 1229–1239 (2003).
- Lamarche, N. *et al.* Production of the R2 subunit of ribonucleotide reductase from herpes simplex virus with prokaryotic and eukaryotic expression systems: higher activity of R2 produced by eukaryotic cells related to higher iron-binding capacity. *Biochem. J.* **320**, 129–135 (1996).
- David, V. *et al.* Identification of cofilin, coronin, Rac and capZ in actin tails using a *Listeria* affinity approach. *J. Cell Sci.* **111**, 2877–2884 (1998).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank L. Blanchoin for help and discussions, L. Machesky for the gift of plasmids Scar-WA and ScarW, and P. Renesto and all members of the Cossart laboratory for discussions and suggestions. This work was supported by the Pasteur Institute, the Ministère de la Recherche et de la Technologie (Programme PRFMMIP) and the Direction Générale des Armées (DGA). C.E. is supported by a Human Frontier Science programme fellowship. P.C. is an international scholar from the Howard Hughes Medical Institute.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.C. (pcossart@pasteur.fr). The sequence for RickA has the GenBank accession number RCO909.

The ADP/ATP translocator is not essential for the mitochondrial permeability transition pore

Jason E. Kokoszka^{1,4*}, Katrina G. Waymire^{1,4}, Shawn E. Levy^{4*}, James E. Sligh^{4*}, Jiyang Cai⁵, Dean P. Jones⁵, Grant R. MacGregor^{1,2,4} & Douglas C. Wallace^{1,3,4}

¹Center for Molecular and Mitochondrial Medicine and Genetics, ²Department of Developmental and Cell Biology, ³Departments of Biological Chemistry and Ecology and Evolutionary Biology, University of California, Irvine, California 92697, USA

⁴Center for Molecular Medicine, ⁵Departments of Medicine and Biochemistry, Emory University School of Medicine, Atlanta, Georgia 30322, USA

* Present addresses: Orchid Cellmark, Germantown, Maryland 20874, USA (J.E.K.); Department of Biomedical Informatics (S.E.L.), and Department of Medicine, Dermatology Division (J.E.S.), Vanderbilt University Medical Center, Nashville, Tennessee 37232, USA

A sudden increase in permeability of the inner mitochondrial membrane, the so-called mitochondrial permeability transition, is a common feature of apoptosis and is mediated by the mitochondrial permeability transition pore (mtPTP). It is thought that the mtPTP is a protein complex formed by the voltage-dependent anion channel, members of the pro- and anti-apoptotic BAX-BCL2 protein family, cyclophilin D, and the adenine nucleotide (ADP/ATP) translocators (ANTs)^{1,2}. The latter exchange mitochondrial ATP for cytosolic ADP and have been implicated in cell death. To investigate the role of the ANTs in the mtPTP, we genetically inactivated the two isoforms of ANT^{3–5} in mouse liver and analysed mtPTP activation in isolated mitochondria and the induction of cell death in hepatocytes. Mito-

chondria lacking ANT could still be induced to undergo permeability transition, resulting in release of cytochrome *c*. However, more Ca^{2+} than usual was required to activate the mtPTP, and the pore could no longer be regulated by ANT ligands. Moreover, hepatocytes without ANT remained competent to respond to various initiators of cell death. Therefore, ANTs are non-essential structural components of the mtPTP, although they do contribute to its regulation.

To investigate the role of ANTs in the mtPTP, we inactivated both the heart-muscle (*Ant1*) and the systemic (*Ant2*) ANT isoform genes in the mouse liver. Humans have three *ANT* genes^{6,7}, whereas results of complementary DNA library screening⁵ and northern and western analyses³ have suggested that mouse has only two *Ant* genes. To verify this, we screened the Celera and Ensembl mouse genome assemblies, as well as the respective EST databases, by BLAST analysis using the *ANT3* cDNA sequence. Consistent with previous results, this produced only two significant matches; that is, *Ant1* and *Ant2*.

Mice with mutations in *Ant1* and *Ant2* were generated by homologous recombination in embryonic stem cells. Mice with a

non-conditional null mutant allele of *Ant1* have been described⁵. A CRE-conditional null mutant allele of the X-linked *Ant2* (*Ant2^{fl}*) was produced by generating a targeting vector in which *loxP* sites flanked exons 3 and 4 (Fig. 1a). Targeted embryonic stem cells

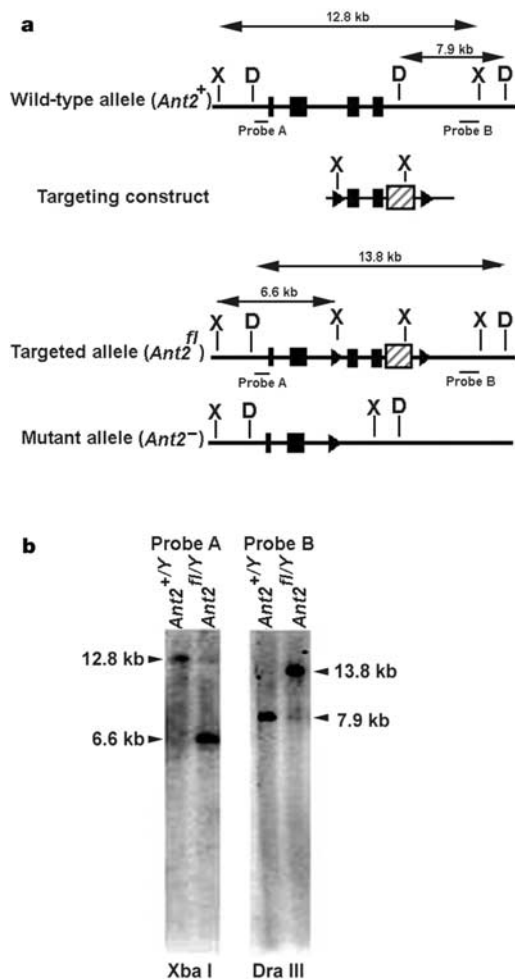


Figure 1 Preparation of a CRE-conditional *Ant2*-null mutant allele in mouse embryonic stem cells. **a**, Top to bottom, maps of the wild-type *Ant2* locus, the region of homology within the targeting construct, and the floxed (*Ant2*^{fl}) and mutant (*Ant2*⁻) alleles. Exons (black boxes), *loxP* sites flanking exons 3 and 4 (triangles), *PGKneo* selection cassette within 3' UTR (hatched box). **a**, **b**, Arrows indicate sizes of Xba I (X) and Dra III (D) restriction fragments used to identify correctly targeted clones by Southern analysis of embryonic-stem-cell genomic DNA using probes A and B. **b**, Southern blot of targeted and untargeted embryonic-stem-cell clones.

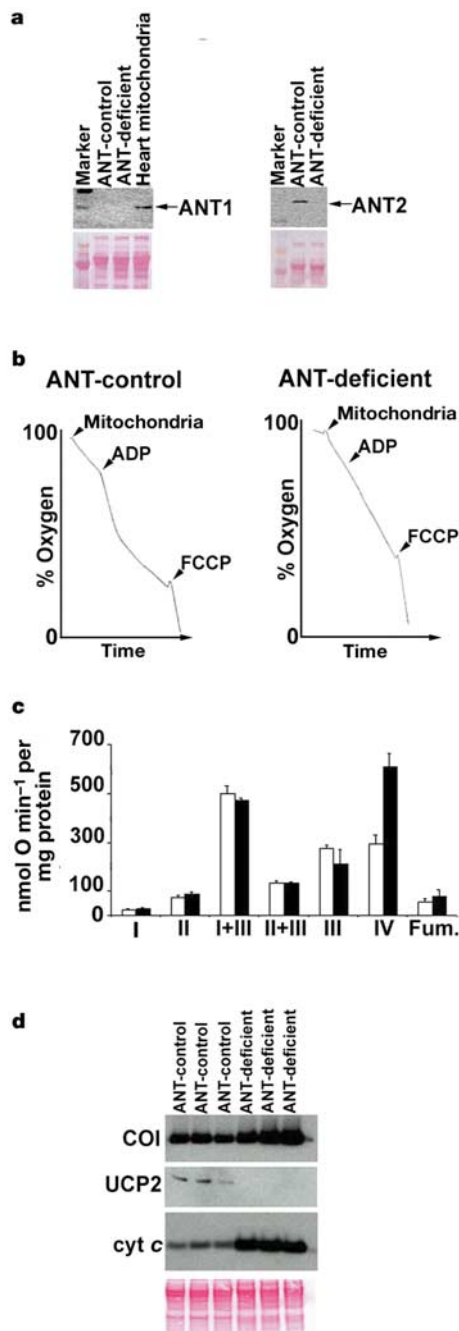


Figure 2 Inactivation of ANT in liver mitochondria. **a**, Western blots of mitochondria using ANT1 and ANT2-specific antibodies. Heart mitochondria are a positive ANT1 control. ANT1 and ANT2 are both 30 kDa (arrows). The Ponceau-S-stained blot (below) is the loading control. **b**, ADP-stimulated mitochondrial respiration. Arrowheads denote addition of 600 μg of mitochondrial protein, 125 nmol of ADP and 75 nM FCCP. **c**, OXPHOS enzyme activities of ANT-control (white bars) or ANT-deficient (black bars) liver mitochondria (mean of three mice each). OXPHOS complexes I–IV and fumarase (Fum.) are shown. Error bars show the standard error. **d**, Western blots of mitochondria from three independent mice for COI (56 kDa), UCP2 (33 kDa), cytochrome (cyt) *c* (11 kDa). The Ponceau S blot (below) is the loading control.

(Fig. 1b) were used to introduce the conditional floxed allele into the mouse germ line. Liver-specific inactivation of the *Ant2^{fl}* allele was achieved by breeding *Ant2^{fl}* mice with an *Alb-Cre* line of transgenic mice that expresses CRE under control of the liver-specific albumin promoter, which results in the excision of exons 3 and 4 of *Ant2* (Fig. 1a). More than 99% of floxed loci are excised in developing hepatocytes of *Alb-Cre* mice⁸. The *Ant2^{fl}* allele was then bred onto an *Ant1^{-/-}* background⁵. The resulting *Ant1^{-/-}, Ant2^{fl/y}* females were crossed with *Ant1^{-/-}, Ant2^{+/-}, Alb-Cre* hemizygous males. Male *Ant1^{-/-}, Ant2^{fl/y}, Alb-Cre* progeny, which are null for both *Ant1* and *Ant2* in their livers, were used to analyse liver mtPTP function. ANT1 is barely detectable in liver³, and *Ant1^{+/-}, Ant2^{fl/y}, Alb-Cre* animals gave the same results as *Ant1^{-/-}, Ant2^{fl/y}, Alb-Cre* mice (data not shown).

The absence of ANT1 and ANT2 protein in liver mitochondria from *Ant1^{-/-}, Ant2^{fl/y}, Alb-Cre* mice was verified by western blot analysis using isoform-specific antibodies (Fig. 2a). Complete ANT deficiency was confirmed by demonstrating that respiration in mitochondria from *Ant1^{-/-}, Ant2^{fl/y}, Alb-Cre* animals could not be stimulated by the addition of ADP (Fig. 2b). Hence, the liver mitochondria of the *Ant1^{-/-}, Ant2^{fl/y}, Alb-Cre* mice have no ANT.

Endogenous respiration rates of ANT1/ANT2-deficient mitochondria were almost twice that of control mitochondria (34.58 ± 1.6 versus 18.12 ± 1.1 nmol O min⁻¹ per mg protein) (Fig. 2b) and the mitochondrial membrane potential ($\Delta P = \Delta \Psi + \Delta pH$) of the ANT-deficient mitochondria was higher than that of controls (191.7 ± 4.9 versus 172.9 ± 3.5 mV). Analysis of the specific activity of OXPHOS enzyme complexes in ANT-deficient mitochondria revealed that complex IV (cytochrome *c* oxidase, COX) was increased more than twofold compared with controls ($P < 0.01$) (Fig. 2c). This was confirmed by western blot analysis, which revealed that the mitochondrial COX subunit I (COI) and cytochrome *c* proteins were more abundant in the ANT-deficient mitochondria (Fig. 2d). Hence, the increased respiration rate is likely to be the result of the specific upregulation in COX activity, suggesting that COX activity may modulate respiration rate. Because ΔP is the product of proton pumping versus proton leak⁹, we analysed the level of the systemic uncoupler protein, UCP2, by western blot. UCP2 was downregulated to undetectable levels in the ANT-deficient mitochondria, suggesting that the increased ΔP resulted from decreased proton leak (Fig. 2d).

We then used the ANT-deficient liver mitochondria to investigate the role of ANT in the structure and regulation of the mtPTP. To determine whether the mtPTP was present in the absence of ANT, we isolated mitochondria from livers of ANT-deficient (*Ant1^{-/-}, Ant2^{fl/y}, Alb-Cre*) and ANT-control (*Ant1^{-/-}, Ant2^{fl/y}*) mice and tested their sensitivity to various inducers of the mtPTP.

Because the mtPTP is a voltage-dependent channel, we determined whether the mtPTP could be activated by the uncoupler carbonyl cyanide *p*-(trifluoromethoxy)phenylhydrazone (FCCP) in the presence of Ca²⁺ (refs 10, 11) (Fig. 3a, b). Opening of the pore was monitored by mitochondrial swelling through decreased light scattering. Strikingly, ANT-deficient and ANT-control mitochondria responded equally well to FCCP-induced mtPTP activation. To confirm that the mitochondrial swelling was due to the activation of the mtPTP, the experiment was repeated in the presence of the mtPTP inhibitor cyclosporin A (CsA). This prevented the swelling of both the ANT-control and ANT-deficient mitochondria (Fig. 3a, b). Hence, the ANT-deficient mitochondria have a functional, CsA-inhibitable mtPTP, demonstrating that ANT is not required to form a functional mtPTP.

To further characterize the mtPTP in ANT-deficient mitochondria, we quantified the propensity of ANT-deficient and ANT-control mitochondria to undergo the permeability transition by determining the concentration of Ca²⁺ required to activate the mtPTP and collapse ΔP . Mitochondria were energized with succinate in the presence of the lipophilic cation tetraphenylphospho-

nium cation (TPP⁺). TPP⁺ is accumulated into the mitochondrial matrix in proportion to ΔP , and its concentration in the suspension medium can be monitored using a TPP⁺-sensitive electrode^{12,13}. Aliquots of Ca²⁺ were added to the reaction, each addition causing a transient release of TPP⁺ as Ca²⁺ was taken into the mitochondria at the expense of ΔP . The Ca²⁺ additions were repeated until sufficient Ca²⁺ accumulated in the mitochondrial matrix to activate the mtPTP, collapsing ΔP and irreversibly releasing the mitochondrial TPP⁺. This process can be inhibited by CsA. Hence, the amount of Ca²⁺ required to cause TPP⁺ release provides a quantitative measurement of the sensitivity of the mtPTP.

ANT-deficient liver mitochondria could still be induced to undergo the CsA-inhibitable permeability transition by sequential addition of Ca²⁺ (Fig. 3c). Thus, the existence of a functional mtPTP in the absence of ANT was confirmed. However, threefold more Ca²⁺ was required to activate the mtPTP in the ANT-deficient mitochondria than in the ANT-control mitochondria (Fig. 3c).

Furthermore, in both ANT-deficient and ANT-control mitochondria, activation of the mtPTP (as monitored by Ca²⁺-induced TPP⁺ release—that is, the permeability transition) was associated with the release of cytochrome *c* from the mitochondria into the supernatant. By contrast, the inner membrane COI protein remained associated with the mitochondrial pellets (Fig. 3d). The ANT-deficient mitochondria contained more cytochrome *c* than did controls, so the mutant mitochondria released more cyto-

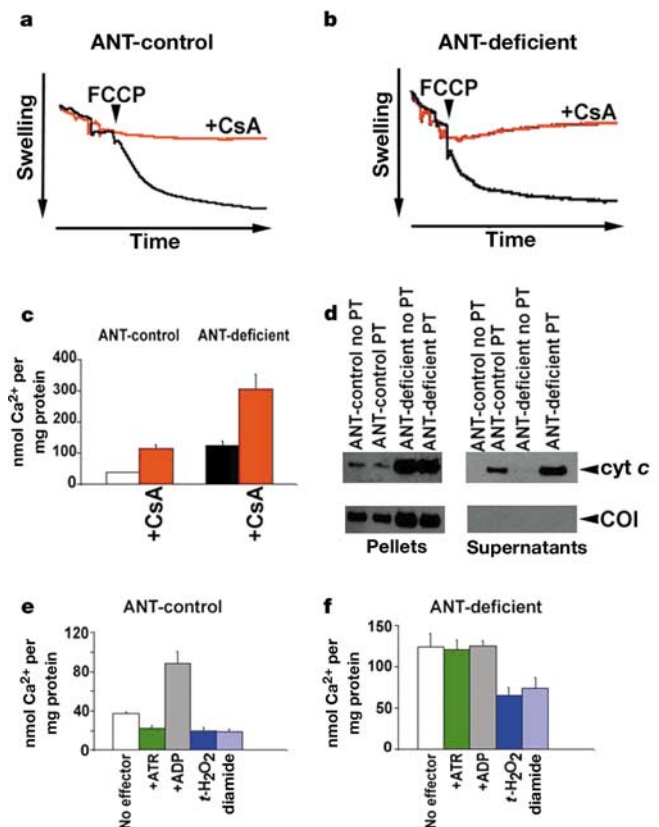


Figure 3 Effect of ANT-deficiency on mtPTP activation in liver mitochondria. **a**, ANT-control and **b**, ANT-deficient mtPTP activation by ΔP depolarization (FCCP) monitored by mitochondrial swelling. Black trace, no CsA; red trace, with CsA. **c**, Amount of Ca²⁺ required per mg of mitochondrial protein to activate the mtPTP and collapse ΔP . **d**, Western blots of cytochrome *c* released into the supernatant fraction versus COI retained in the mitochondrial pellet following Ca²⁺-activation of the mtPTP. **e**, ANT-control and **f**, ANT-deficient Ca²⁺-activation of the mtPTP in the presence of positive and negative effectors.

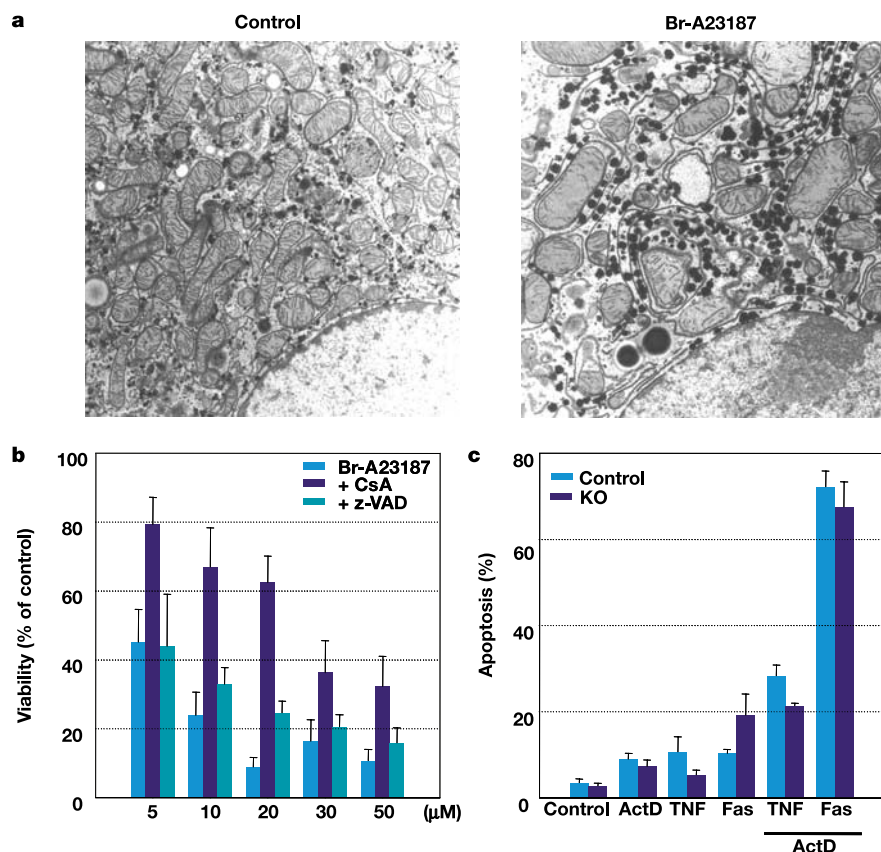


Figure 4 Induction of cell death in ANT-deficient mouse hepatocytes. **a**, Transmission electron microscope images of ANT-deficient hepatocytes. Control, untreated; Br-A23187, treated for 1 h with 10 μ M Br-A23187. **b**, Calcium ionophore (Br-A23187) induced cell death (% of LDH released) of ANT-deficient treated hepatocytes relative to

untreated hepatocytes, without or with pretreatment by either CsA or z-VAD. **c**, Receptor-mediated apoptosis (Hoechst 33258 analysed nuclear fragmentation) induced by TNF- α or soluble Fas ligand treatment, with or without actinomycin D (ActD). KO, knockout.

chrome *c* (Fig. 3d). Hence, ANT₂s are also not required for mtPTP-induced cytochrome *c* release, confirming that the mtPTP can function in the absence of ANT.

Next, we investigated whether ANT₂s regulate the function of the mtPTP. ANT-deficient and control liver mitochondria were loaded with TPP⁺ and then exposed to various mtPTP effectors. *Tert*-butyl hydroperoxide (*t*-H₂O₂), a non-specific oxidant, and diamide, a specific -SH group oxidant, are both inducers of the mtPTP¹⁴. These compounds increased the predilection of the mtPTP to undergo the permeability transition, independent of the presence of ANT (Fig. 3e, f). The ANT ligands atractyloside (ATR) and ADP are positive and negative effectors of the mtPTP, respectively^{15,16}. Although ATR activated and ADP inhibited the mtPTP in ANT-control mitochondria, they had no effect on ANT-deficient mitochondria (Fig. 3e, f). Thus, the absence of ANT results in the loss of the modulation of the mtPTP by ANT ligands, including adenine nucleotides.

We next investigated whether absence of ANT₂s affected the response of cultured hepatocytes to inducers of apoptosis and necrosis (Fig. 4). Mitochondrial swelling (Fig. 4a) and cell death (Fig. 4b) occurred in ANT-deficient (*Ant1*^{-/-}, *Ant2*^{fl/y}, *Alb-Cre*) hepatocytes after exposure to the calcium ionophore Br-A23187 (ref. 17). This was partially inhibited by pretreatment with CsA but not with the general caspase inhibitor z-VAD (Fig. 4b). In fact, ANT-deficient hepatocytes were more sensitive to Br-A23187 than were ANT-control cells, suggesting that ANT-deficiency may pre-sensitize the mitochondria to the permeability transition. Moreover, no difference was observed in the sensitivity of ANT-deficient and ANT-control hepatocytes to treatment with soluble Fas ligand or

tumour-necrosis factor- α (TNF- α) in the presence of actinomycin D¹⁸ (Fig. 4c). Hence, ANT₂s are not required for mitochondrial permeability transition or for these forms of mitochondria-mediated hepatocyte cell death.

In conclusion, our studies reveal that the ANT₂s are non-essential components of the mtPTP and that they are dispensable for at least some forms of mtPTP-associated cell death. However, the ANT₂s do have an essential role in regulating permeability transition by modulating the sensitivity of the mtPTP to Ca²⁺ activation and ANT ligands. □

Methods

Ant2^{fl} mice on a mixed 129S4 and C57BL/6 background were generated using standard methods under a protocol approved by Emory University's IACUC. The targeting vector contained a floxed *Ant2* gene with a PGK-*neo* cassette added to the 3'-untranslated region (Fig. 1a). Targeted AK7.1 embryonic stem cell clones were identified by Southern analysis (Fig. 1b). The 5'-probe (probe A) detected a 12.8-kilobase (kb) *Xba* I fragment for the wild-type *Ant2* allele and a 6.6-kb fragment for the targeted allele. The 3'-probe (probe B) detected a 7.9-kb *Dra* III fragment for the wild-type allele and a 13.8-kb fragment for the targeted allele. Site-specific recombination between the *loxP* sites removed the last third of the *Ant2* coding sequence, including putative transmembrane domains 5 and 6, plus the PGK-*neo* cassette, generating a non-functional protein.

Ant2 was genotyped by PCR (forward primer 5'-ACTCAACCTAGGGCCCTTG-3', reverse primer 5'-GGGAGCATTCCTGAAAAATAA-3'; 35 cycles at 94 °C for 20 s; at 56 °C for 30 s; at 72 °C for 40 s) to detect the targeted (485 base pairs, bp) and wild-type (384 bp) alleles. The mutant *Ant2* allele (850 bp) was detected using the same forward primer and reverse primer 5'-GACTTACCCTCCACGACAGC-3'; 35 cycles at 94 °C for 20 s; at 65 °C for 30 s; at 72 °C for 60 s. Genotyping at *Ant1* was determined as described⁵.

For western blots, isoform-specific ANT1 and ANT2 antibodies⁵, and antibodies recognizing COI, UCP2 and cytochrome *c* (Molecular Probes and Zymed) were employed. These were reacted to isolated mitochondrial protein (20 μ g) or supernatants separated by SDS-PAGE and blotted onto nitrocellulose.

Mitochondria were isolated from liver by homogenization followed by differential

centrifugation. Respiration and OXPHOS enzyme activities were normalized for protein concentration using the Coomassie Stain kit (Pierce)^{13,19}. Mitochondrial membrane potential was calculated from mitochondrial uptake of TPP⁺ using a TPP⁺-sensitive electrode¹². The sensitivity of the mtPTP to undergo permeability transition was examined in liver mitochondria of 10-month-old *Ant1*^{-/-}, *Ant2*^{fl/y}, *Alb-Cre* animals by TPP⁺ release after sequential additions of 10 nM CaCl₂ (ref. 13). The mtPTP modulators used were 1 mM t-H₂O₂, 0.1 mM diamide, 100 μM ATR and 125 μM ADP (Fig. 3).

Mitochondrial PTP activation was also monitored by mitochondrial swelling using light scattering at 546 nm for 10 min in 1.5 ml with 1 mg mitochondrial protein and 16.5 nmol CaCl₂. The reaction was initiated by the addition of 0.1 μM ruthenium red and 1 μM FCCP.

Hepatocytes were isolated from 12- to 15-month-old anaesthetized mice, perfused *in situ* with collagenase-dispase medium (Invitrogen). Hepatocytes were gently released, filtered and cultured on collagen-coated cover glasses or plates in Waymouth's MB-752/1 medium¹⁷. Hepatocyte sensitivity to Ca²⁺ ionophore was examined by extent of cell death assessed by the percentage of total cellular lactate dehydrogenase (LDH) released into the medium²⁰ after treatment with 5 to 50 μM Br-A23187 for 1 h, without or with a 30 min pretreatment of either 1 μM CsA or 50 μM z-VAD 17 (Fig. 4). Receptor-induced cell death was monitored by analysis of nuclear morphology using Hoechst 33258 (Molecular Probes) staining after treatment with 100 ng ml⁻¹ of murine recombinant TNF-α (R&D Systems) or 4 ng ml⁻¹ of human recombinant Fas ligand (Upstate), with or without 0.2 μg ml⁻¹ actinomycin D (Fig. 4).

Received 14 August; accepted 10 November 2003; doi:10.1038/nature02229.

- Zoratti, M. & Szabo, I. The mitochondrial permeability transition. *Biochim. Biophys. Acta* **1241**, 139–176 (1995).
- Marzo, I. *et al.* Bax and adenine nucleotide translocator cooperate in the mitochondrial control of apoptosis. *Science* **281**, 2027–2031 (1998).
- Levy, S. E., Chen, Y.-S., Graham, B. H. & Wallace, D. C. Expression and sequence analysis of the mouse adenine nucleotide translocase 1 and 2 genes. *Gene* **254**, 57–66 (2000).
- Ellison, J. W., Salido, E. C. & Shapiro, L. J. Genetic mapping of the adenine nucleotide translocase-2 gene (*Ant2*) to the mouse proximal X chromosome. *Genomics* **36**, 369–371 (1996).
- Graham, B. *et al.* A mouse model for mitochondrial myopathy and cardiomyopathy resulting from a deficiency in the heart/skeletal muscle isoform of the adenine nucleotide translocator. *Nature Genet.* **16**, 226–234 (1997).
- Stepien, G., Torroni, A., Chung, A. B., Hodge, J. A. & Wallace, D. C. Differential expression of adenine nucleotide translocator isoforms in mammalian tissues and during muscle cell differentiation. *J. Biol. Chem.* **267**, 14592–14597 (1992).
- Lunardi, J., Hurko, O., Engel, W. K. & Attardi, G. The multiple ADP/ATP translocase genes are differentially expressed during human muscle development. *J. Biol. Chem.* **267**, 15267–15270 (1992).
- Postic, C. & Magnuson, M. A. DNA excision in liver by an albumin-Cre transgene occurs progressively with age. *Genesis* **26**, 149–150 (2000).
- Boss, O., Hagen, T. & Lowell, B. B. Uncoupling proteins 2 and 3: potential regulators of mitochondrial energy metabolism. *Diabetes* **49**, 143–156 (2000).
- Petronilli, V., Niccoli, A., Costantini, P., Colonna, R. & Bernardi, P. Regulation of the permeability transition pore, a voltage-dependent mitochondrial channel inhibited by cyclosporin A. *Biochim. Biophys. Acta* **1187**, 255–259 (1994).
- Bernardi, P. Modulation of the mitochondrial cyclosporin A-sensitive permeability transition pore by the proton electrochemical gradient. Evidence that the pore can be opened by membrane depolarization. *J. Biol. Chem.* **267**, 8834–8839 (1992).
- Esposito, L. A. *et al.* Mitochondrial oxidative stress in mice lacking the glutathione peroxidase-1 gene. *Free Radic. Biol. Med.* **28**, 754–766 (2000).
- Kokoszka, J. E., Coskun, P., Esposito, L. & Wallace, D. C. Increased mitochondrial oxidative stress in the *Sod2*^{-/-} mouse results in the age-related decline of mitochondrial function culminating in increased apoptosis. *Proc. Natl Acad. Sci. USA* **98**, 2278–2283 (2001).
- Halestrap, A. P., Woodfield, K. Y. & Connern, C. P. Oxidative stress, thiol reagents, and membrane potential modulate the mitochondrial permeability transition by affecting nucleotide binding to the adenine nucleotide translocase. *J. Biol. Chem.* **272**, 3346–3354 (1997).
- Lapidus, R. G. & Sokolove, P. M. The mitochondrial permeability transition. Interactions of spermine, ADP, and inorganic phosphate. *J. Biol. Chem.* **269**, 18931–18936 (1994).
- Novgorodov, S. A., Gud, T. I., Brierley, G. P. & Pfeiffer, D. R. Magnesium ion modulates the sensitivity of the mitochondrial permeability transition pore to cyclosporin A and ADP. *Arch. Biochem. Biophys.* **311**, 219–228 (1994).
- Qian, T., Herman, B. & Lemasters, J. J. The mitochondrial permeability transition mediates both necrotic and apoptotic death of hepatocytes exposed to Br-A23187. *Toxicol. Appl. Pharmacol.* **154**, 117–125 (1999).
- Hatano, E. *et al.* The mitochondrial permeability transition augments Fas-induced apoptosis in mouse hepatocytes. *J. Biol. Chem.* **275**, 11814–11823 (2000).
- Trounce, I. A., Kim, Y. L., Jun, A. S. & Wallace, D. C. Assessment of mitochondrial oxidative phosphorylation in patient muscle biopsies, lymphoblasts, and transmembrane cell lines. *Methods Enzymol.* **264**, 484–509 (1996).
- Leist, M. *et al.* Murine hepatocyte apoptosis induced *in vitro* and *in vivo* by TNF-α requires transcriptional arrest. *J. Immunol.* **153**, 1778–1788 (1994).

Acknowledgements We thank M. Magnuson for providing the *Alb-Cre* transgenic mice, L. Hayes for mouse husbandry and genotyping, and H. Yi for the electron microscope analysis. This work was funded by US National Institutes of Health grants awarded to D.C.W., G.R.M. and D.P.J.

Competing interests statement The authors declare competing financial interests: details accompany the paper on www.nature.com/nature.

Correspondence and requests for materials should be addressed to D.C.W. (dwallace@uci.edu).

Mechanically driven ATP synthesis by F₁-ATPase

Hiroyasu Itoh^{1,2}, Akira Takahashi³, Kengo Adachi⁴, Hiroyuki Noji⁵, Ryohei Yasuda⁶, Masasuke Yoshida⁷ & Kazuhiko Kinoshita Jr⁴

¹Tsukuba Research Laboratory, Hamamatsu Photonics KK, and ²CREST “Creation and application of soft nano-machine, the hyperfunctional molecular machine” Team 13*, Tokodai, Tsukuba 300-2635, Japan

³System Division, Hamamatsu Photonics KK, Joka, Hamamatsu 431-3103, Japan

⁴Center for Integrative Bioscience, Okazaki National Research Institutes, Okazaki 444-8585, Japan

⁵Institute of Industrial Science, University of Tokyo, Tokyo 153-8505, Japan

⁶Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724 USA

⁷ERATO “ATP System”, 5800-3 Nagatsuta, Yokohama 226-0026, Japan

ATP, the main biological energy currency, is synthesized from ADP and inorganic phosphate by ATP synthase in an energy-requiring reaction^{1–3}. The F₁ portion of ATP synthase, also known as F₁-ATPase, functions as a rotary molecular motor: *in vitro* its γ-subunit rotates⁴ against the surrounding α₃β₃ subunits⁵, hydrolysing ATP in three separate catalytic sites on the β-subunits. It is widely believed that reverse rotation of the γ-subunit, driven by proton flow through the associated F₀ portion of ATP synthase, leads to ATP synthesis in biological systems^{1–3,6,7}. Here we present direct evidence for the chemical synthesis of ATP driven by mechanical energy. We attached a magnetic bead to the γ-subunit of isolated F₁ on a glass surface, and rotated the bead using electrical magnets. Rotation in the appropriate direction resulted in the appearance of ATP in the medium as detected by the luciferase–luciferin reaction. This shows that a vectorial force (torque) working at one particular point on a protein machine can influence a chemical reaction occurring in physically remote catalytic sites, driving the reaction far from equilibrium.

When isolated F₁ hydrolyses ATP, its central γ-subunit rotates anticlockwise⁴ when viewed from above in Fig. 1a, with an efficiency of chemical-to-mechanical energy conversion approaching 100% (ref. 8). The purpose of this study was to show that the chemo-mechanical coupling in the F₁ motor is completely reversible, and that reversal is achieved by manipulating a single variable—that is, the rotary angle of the γ-subunit. Any molecular machine would be reversible if one could manipulate all constituent atoms at will. Whether one or a few thermodynamic handles exist in a chemo-mechanical molecular machine such that its operation can be controlled through that handle in both directions is an important but unresolved issue. For example, whether one can synthesize ATP by pulling back a linear molecular motor such as myosin or kinesin—and if so, where to pull—is unknown. Reversal of the whole ATP synthase is well documented^{1,9}, including the demonstration of γ-subunit reorientation under synthesis conditions¹⁰, but whether or not the γ-subunit angle serves as a single handle for F₁ reversal has not been tested.

To prove this reversibility, we used α₃β₃γ, the minimal subcomplex of F₁ that shows ATP-catalysed rotation⁴. The subcomplex was attached to a glass surface through histidine residues engineered at the amino terminus of the β-subunits, and a magnetic bead coated with streptavidin was attached to the γ-subunit, which had been biotinylated at two engineered cysteines (Fig. 1a). The beads were rotated with magnets (Fig. 1b–d) in a medium containing ADP and phosphate as substrates and the luciferin–luciferase system^{11,12}, which emits a photon when it captures and hydrolyses ATP. The initial idea was to count these chemiluminescent photons (Fig. 1b); however, background luminescence originating from contaminant ATP present in ADP even after purification was a problem. Thus, the

volume of medium per active F_1 molecule had to be small.

First, we tried to reduce the volume by making microdroplets in oil (Fig. 2a). Figure 2b shows data from a 4×4 array of droplets in one chamber. The beads in droplets were rotated at 10 Hz alternately for 5 min each in the direction of hydrolysis (anticlockwise when viewed from top in Fig. 1a) and synthesis (clockwise). As seen in Fig. 2b, 14 out of 16 droplet curves showed the M-shaped pattern of photon counts expected for the sequence of rotation direction when the overall decline was taken into account. The decline was due to the gradual disappearance of the aqueous phase into oil: although we saturated the oil with water before the experiment, droplets tended to shrink over time. For random photon counts, the probability of observing a slanting M shape is 8^{-1} . The probability of observing 14 or more M shapes out of 16 is 2×10^{-11} . The data set shown in Fig. 2b thus strongly indicates mechanical synthesis. The experiment is extremely difficult (at most a few beads rotate in each droplet), however, and we have obtained only a few more data sets that contained several M-shaped patterns.

We thus tried to increase the number of rotating beads in an ordinary observation chamber (Fig. 1c) by infusing a concentrated solution of beads carrying F_1 . To allow F_1 to rotate in the proper direction, we derivatized only the bottom surface of the chamber with nickel nitrilo-triacetic acid (Ni^{2+} -NTA), which would specifically bind the β -histidines. In control experiments done in 4 mM ATP (no ADP) and without magnets, we found in the field of view of $1.0 \times 10^5 \mu m^2$ as many as 480 ± 70 F_1 molecules rotating anticlockwise at the bottom (three chambers). However, the high density of beads resulted in nonspecific binding to the ceiling, where 100 ± 20 beads rotated clockwise (as viewed from above the chamber). We also tested in the ATP medium whether forced rotation by external magnets would damage F_1 . After confirming ATP-driven rotation, we turned on the magnets and applied several bursts of hundreds of revolutions at 10 Hz in both directions. When

the magnets were turned off, ATP-driven rotation resumed.

Synthesis was shown in the ADP-luciferase medium by accumulating chemiluminescence photons over a series of 5-min intervals in which the magnetic field was rotated at 10 Hz in either direction or turned off. All series produced the expected pattern (Fig. 3a): higher photon counts during clockwise rotation (S, synthesis) than during anticlockwise rotation (H, hydrolysis) or no rotation (N). This graph compiles all data taken in consecutive experiments (about half of the experiments failed at some point, for example during chamber preparation or because of a large focus drift, and did not produce data). Mechanical synthesis in the flat chamber was reproducible, although variation among data was still large.

We note that in most curves shown in Fig. 3a including the total counts, counts during anticlockwise rotation (H) are higher than those at no rotation (N). This is due to the presence of F_1 at the ceiling of the chamber as stated above. For these upside-down F_1 molecules, anticlockwise bead rotation will drive ATP synthesis. Indeed, when we flipped the chamber upside down after obtaining the unbroken blue line in Fig. 3a, the count pattern was reversed, as shown by the broken blue line. Another reason for the higher counts during anticlockwise rotation was that luciferase did not consume all of the newly synthesized ATP in 5 min: as shown by the unbroken lines in Fig. 3b, the luminescence at no rotation was high after ATP synthesis at the arrows. Luminescence decay after ATP mixing was shown to involve a component with a lifetime of about 3 min (see Supplementary Information). Taking all of the above points into account, we consider that mechanical synthesis has been conclusively demonstrated.

Figure 4 shows the effect of rotary speed on the efficiency of ATP synthesis. The synthesis rate apparently saturated above 3 Hz (Fig. 4b). This was because larger beads or bead aggregates failed to rotate at high speeds, as confirmed by direct observation (uncoupling between magnet and bead rotations). Calibration of the

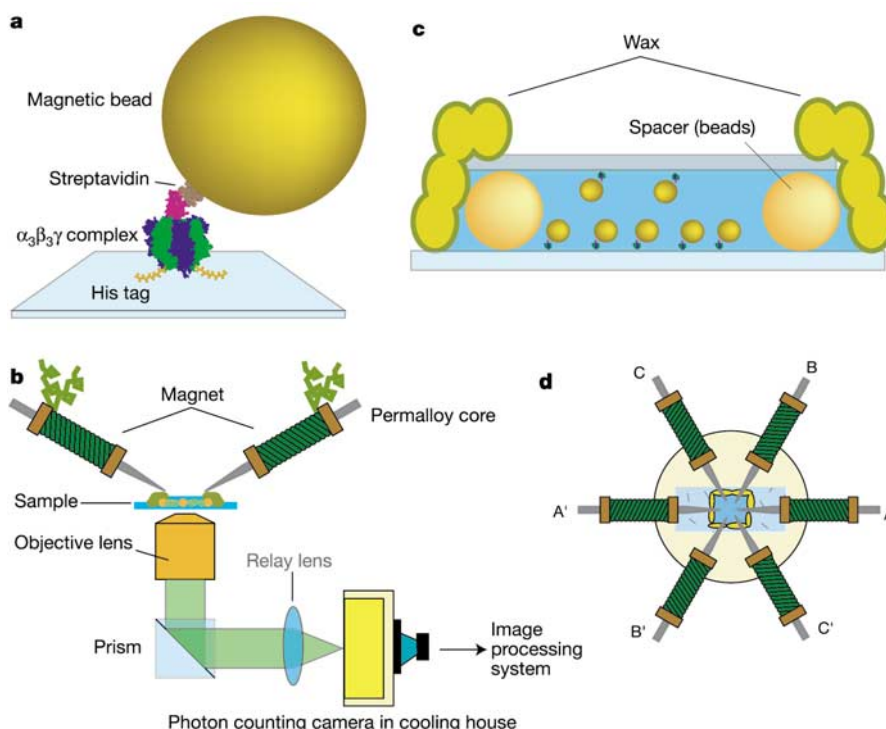


Figure 1 Experimental set-up. **a**, Basic design. The structures of F_1 and streptavidin are from ref. 21 and ref. 22, respectively. The bead is not to scale and the orientation of streptavidin is unknown. **b**, Side view of the optical system. **c**, Observation chamber.

Ni^{2+} -NTA was applied only to the bottom surface, but some F_1 -conjugated beads rotated on the ceiling. **d**, Top view of the magnets.

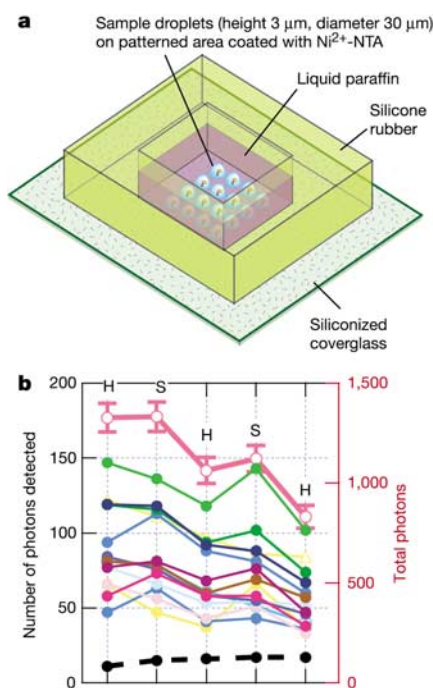


Figure 2 Rotational synthesis in microdroplets. **a**, Observation chamber. On the siliconized bottom coverglass, 4×4 spots separated by $\sim 50 \mu\text{m}$ and measuring $20\text{--}30 \mu\text{m}$ were derivatized with Ni^{2+} -NTA, which rendered the spots hydrophilic. An oil layer with a thickness of $\sim 3 \text{ mm}$ was placed on top. Using a glass micropipette, we formed a droplet of $\sim 1 \text{ pl}$ containing F_1 on each spot and then replaced the solution with one containing beads, ADP, phosphate and the chemiluminescence system. **b**, Simultaneous observation of 16 droplets. Thick magenta curve with open circles represents the sum of all unbroken curves; error bars represent $\pm 2\sigma$, where σ is the expected s.d. for photon statistics (square root of the total count). The broken curve at the bottom represents a control in which an area outside the droplets was imaged. See Supplementary Information for details.

photon-to-ATP ratio (see Supplementary Information) indicated a synthesis rate of about five molecules of ATP per second under rotation at 3 Hz, as compared with the nine molecules of ATP per second expected for the one ATP molecule per 120° scheme⁸.

Although Fig. 4 represents our best data so far and variation among chambers is large (Fig. 3a), we anticipate that the coupling between mechanical rotation of the γ -subunit and chemical synthesis is tight, at least at low speeds. The significant synthesis at low speeds (Fig. 4), which were much lower than the maximal rotary speed of 130 Hz of this motor during ATP hydrolysis¹³, merits attention. Because the slow rotation was at a constant speed, it is likely that chemical reactions were at quasi-equilibrium at all angles. Rotation by ATP hydrolysis can also be made slow and at a constant speed by attaching a long actin filament to the γ -subunit⁸, again suggesting quasi-equilibrium. The implication is that ATP synthesis in F_1 proceeds as a straightforward reversal of the hydrolysis reaction, tracking the same reaction pathway in the opposite direction. On that pathway, both hydrolysis and synthesis reactions are controlled by one mechanical handle, the rotary angle of the γ -subunit. The situation contrasts with the more complex rotary motor of bacterial flagella, where rotational directions can be switched without reversing the proton motive force¹⁴.

ATP synthesis is a chemical reaction that is energetically uphill, requiring $80\text{--}100 \text{ pN nm}$ of energy under physiological conditions². In the experiments shown here, contaminant ATP amounted to about 1 nM, implying that roughly 30 pN nm of free energy was needed per molecule of ATP synthesized (see Supplementary

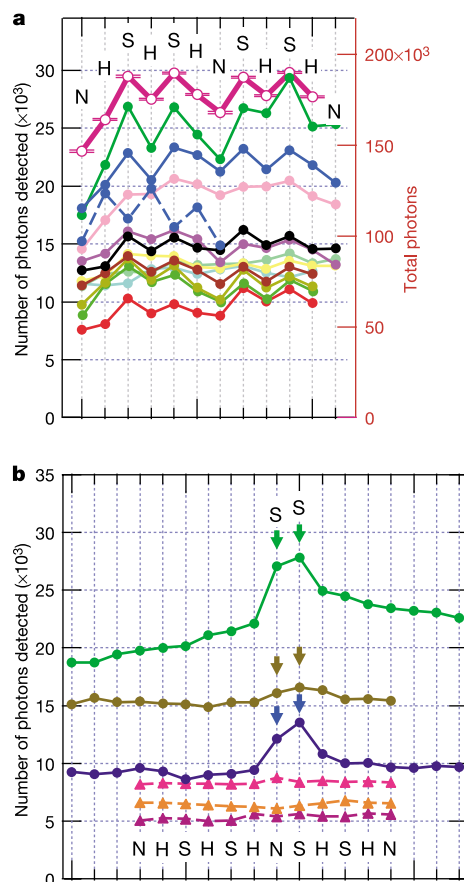


Figure 3 Rotational synthesis in a flat chamber. Each symbol shows the number of photons detected over 5 min in an image area of $7.8 \times 10^4 \mu\text{m}^2$. N, no rotation; H, 10-Hz rotation in the hydrolysis direction (for F_1 on the bottom of the chamber); S, 10-Hz rotation in the synthesis direction. **a**, Results from 12 chambers, distinguished by different colours. Broken blue curve represents results obtained after the chamber giving rise to the unbroken blue curve was flipped upside down. Thick magenta curve with open circles represents the sum of all unbroken curves; error bars represent $\pm 4\sigma$ ($>99.99\%$ confidence). For clarity, some curves have been shifted vertically within $\pm 1,000$ counts. **b**, Control experiments. Unbroken curves show results either without rotation or with imposed clockwise rotation for synthesis at 10 Hz (arrows). Broken curves show experiments carried out as in **a** except that phosphate was omitted from the medium. The counts in the broken curves are low in comparison to the others because phosphate tended to increase the background, apparently by increasing the portion of luciferase that reacted with ATP extremely slowly (see Supplementary Information).

Information). If F_1 , or the motor enzyme myosin, is mixed with high concentrations of ADP and phosphate in the absence of ATP, some ATP is spontaneously formed on the enzyme without input of energy^{15–18}. This, however, is a dead-end reaction and the ATP that has been formed is not available in the medium: release of the tightly bound ATP requires an external supply of energy^{1,13}.

Here we have demonstrated repetitive synthesis by F_1 ($\sim 10^3$ ATP molecules in 5 min), leading to appearance of the product in the medium. To our knowledge, this is the first accomplishment of artificial chemical synthesis by a vectorial force (although nature presumably has been doing this for millions of years). Pressure could also shift a chemical equilibrium by acting on substrates (and solvent); however, in our experiments the chemical equilibrium *per se* is on the side of almost complete hydrolysis, and a force on a point remote from substrates counters the hydrolysis reaction and pushes the equilibrium to the point of favouring synthesis. Because we still have to rely on nature's nanomachine, the F_1 motor, our

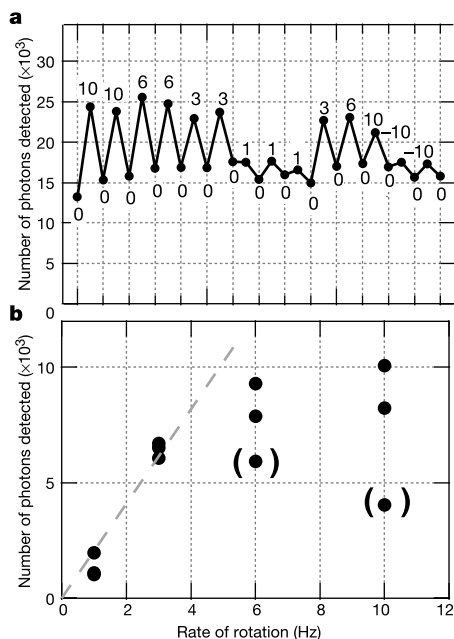


Figure 4 Dependence of synthesis efficiency on the rate of magnet rotation. **a**, Photons detected over an image area of $1.0 \times 10^5 \mu\text{m}^2$ in 5-min periods at various rotary speeds (shown in Hz; negative numbers indicate rotation in the hydrolysis direction). **b**, Photon increments during rotation. From the photon count for each synthesis rotation in **a**, the average of the adjacent no-rotation counts was subtracted. Symbols in the parentheses show the last two measurements for synthesis in **a**; low values for these presumably reflect sample deterioration.

primary goal is to understand fully how it works and thereby to exploit the mechanism in artificial ways. Improving the present system for more quantitative assays, such as the torque and speed dependence of the coupling efficiency, that can be compared with theoretical predictions^{7,19} should be our next task. The key is to obtain magnetic beads that are small and uniform in size. □

Methods

Materials

A mutant $\alpha_3\beta_3\gamma$ subcomplex (comprising C193S α -, His₁₀- β -, and S107C, I210C γ -subunits; referred to as F₁ in this paper), derived from the thermophilic *Bacillus* strain PS3, was biotinylated at the only cysteines on the γ -subunit (ref. 8). Streptavidin-coated magnetic beads (Seradyn; nominally 0.7 μm) were lightly centrifuged to remove large beads and aggregates (but elimination was incomplete). Biotinylated F₁ (400 pM) was incubated with beads (~50 pM) in buffer A (50 mM 3-(*N*-morpholino)propanesulphonic acid/KOH, 20 mM K₂SO₄, 4 mM MgSO₄, pH 7.6) for 10 min at 23 °C, washed with buffer A, and concentrated with a magnet. Luciferase¹² was a kind gift from Kikkoman Co. We purified ADP (K-salt; Sigma) as described²⁰ on a Poros HQ/L column (Applied Biosystems).

Observation chamber

For the chamber shown in Fig. 1c, a 32 × 24 mm² coverglass was functionalized by a silane coupling agent with an SH group (TSL 8380; GE Toshiba Silicone), and reacted first with 10 mg ml⁻¹ maleimide-C₃-NTA (Dojindo) and then with 10 mM NiCl₂. Two parallel strips of Lumirror polyester film (TORAY) were placed on the coverglass, and a 18 × 18 mm² coverglass coated with hexamethyldisilazane was placed on top to form a flow chamber. We infused ~500 pM F₁-conjugated magnetic beads in buffer A₃ (buffer A containing 3 mg ml⁻¹ bovine serum albumin (BSA)) and incubated the chamber for 30 min at 23 °C. After infusing buffer A₃ several times, buffer A containing 10 μM luciferase, 1 mM luciferin, 10 mM K₂PO₄, 200 μM ADP and 1.5 mg ml⁻¹ BSA, together with a small number of 3- μm Dynabeads M-280 beads (Dyna) that would eventually serve as spacers, was infused. The spacer strips were carefully removed and the upper coverglass was pressed to reduce the chamber height to ~3 μm , as determined by the spacer beads. The chamber was then sealed with wax and subjected to observation.

Microscopy

The chamber was placed on an inverted ICM 450 microscope (Zeiss). Luminescence was collected with an oil immersion ×60 objective, numerical aperture 1.45 (Olympus) and deflected with a prism to a factory-made side port (Fig. 1b). The beam was focused with an ED Plan ×2 objective (Nikon) onto a cooled photon-counting camera (V8070U-64-N230/C4566 equipped with an Argus 50 image processing system; Hamamatsu Photonics), which recorded centroid positions of incoming photons (counts in the dark: ~15 s⁻¹ over the whole image plane). To rotate the beads, we placed three opposing pairs of custom-made electromagnets with a Permalloy core on the chamber (Fig. 1d). The three pairs were activated with a custom circuit 120° out of phase to produce a rotating magnetic field (in either direction).

Received 5 May; accepted 31 October 2003; doi:10.1038/nature02212.

- Boyer, P. D. The ATP synthase—a splendid molecular machine. *Annu. Rev. Biochem.* **66**, 717–749 (1997).
- Kinosita, K. Jr, Yasuda, R., Noji, H. & Adachi, K. A rotary molecular motor that can work at near 100% efficiency. *Phil. Trans. R. Soc. Lond. B* **355**, 473–489 (2000).
- Yoshida, M., Muneyuki, E. & Hisabori, T. ATP synthase—a marvellous rotary engine of the cell. *Nature Rev. Mol. Cell Biol.* **2**, 669–677 (2001).
- Noji, H., Yasuda, R., Yoshida, M. & Kinosita, K. Jr Direct observation of the rotation of F₁-ATPase. *Nature* **386**, 299–302 (1997).
- Abrahams, J. P., Leslie, A. G. W., Lutter, R. & Walker, J. E. Structure at 2.8 Å resolution of F₁-ATPase from bovine heart mitochondria. *Nature* **370**, 621–628 (1994).
- Boyer, P. D. & Kohlbrenner, W. E. in *Energy Coupling in Photosynthesis* (eds Selman, B. R. & Selman-Reimer, S.) 231–240 (Elsevier, Amsterdam, 1981).
- Oosawa, F. & Hayashi, S. The loose coupling mechanism in molecular machines of living cells. *Adv. Biophys.* **22**, 151–183 (1986).
- Yasuda, R., Noji, H., Kinosita, K. Jr & Yoshida, M. F₁-ATPase is a highly efficient molecular motor that rotates with discrete 120° steps. *Cell* **93**, 1117–1124 (1998).
- Turina, P., Samoray, D. & Gräber, P. H⁺/ATP ratio of proton transport-coupled ATP synthesis and hydrolysis catalysed by CF₀F₁-liposomes. *EMBO J.* **22**, 418–426 (2003).
- Zhou, Y., Duncan, T. M. & Cross, R. L. Subunit rotation in *Escherichia coli* F₀F₁-ATP synthase during oxidative phosphorylation. *Proc. Natl Acad. Sci. USA* **94**, 10583–10587 (1997).
- McElroy, W. D., Seliger, H. H. & White, E. H. Mechanism of bioluminescence, chemiluminescence and enzyme function in the oxidation of firefly luciferin. *Photochem. Photobiol.* **10**, 153–170 (1969).
- Hattori, N., Kajiyama, N., Maeda, M. & Murakami, S. Mutant luciferase enzymes from fireflies with increased resistance to benzalkonium chloride. *Biosci. Biotechnol. Biochem.* **66**, 2587–2593 (2002).
- Yasuda, R., Noji, H., Yoshida, M., Kinosita, K. Jr & Itoh, H. Resolution of distinct rotational substeps by submillisecond kinetic analysis of F₁-ATPase. *Nature* **410**, 898–904 (2001).
- Berg, H. C. The rotary motor of bacterial flagella. *Annu. Rev. Biochem.* **72**, 19–54 (2003).
- Yoshida, M. The synthesis of enzyme-bound ATP by the F₁-ATPase from the thermophilic bacterium PS3 in 50% dimethylsulfoxide. *Biochem. Biophys. Res. Commun.* **114**, 907–912 (1983).
- Sakamoto, J. Effect of dimethylsulfoxide on ATP synthesis by mitochondrial soluble F₁-ATPase. *J. Biochem.* **96**, 483–487 (1984).
- Wolcott, R. G. & Boyer, P. D. The reversal of the myosin and actomyosin ATPase reactions and the free energy of ATP binding to myosin. *Biochem. Biophys. Res. Commun.* **57**, 709–716 (1974).
- Mannherz, H. G., Schenck, H. & Goody, R. S. Synthesis of ATP from ADP and inorganic phosphate at the myosin-subfragment 1 active site. *Eur. J. Biochem.* **48**, 287–295 (1974).
- Wang, H. & Oster, G. Energy transduction in the F₁ motor of ATP synthase. *Nature* **396**, 279–282 (1998).
- Oishi, N. & Sugi, H. *In vitro* ATP-dependent F-actin sliding on myosin is not influenced by substitution or removal of bound nucleotide. *Biochim. Biophys. Acta* **1185**, 346–349 (1994).
- Menz, R. L., Walker, J. E. & Leslie, A. G. W. Structure of bovine mitochondrial F₁-ATPase with nucleotide bound to all three catalytic sites: implications for the mechanism of rotary catalysis. *Cell* **106**, 331–341 (2001).
- Freitag, S., Trong, I. L., Klumb, L., Stayton, P. S. & Stenkamp, R. E. Structural studies of the streptavidin binding loop. *Protein Sci.* **6**, 1157–1166 (1997).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank T. Hayakawa and T. Hiruma of Hamamatsu Photonics KK who allowed H.I. to work on this project for more than 6 years; S. Brenner for the idea of using microdroplets; M. Sugai for initial work; M. Shio, members of the former CREST Team 13 and the current Kinosita and Yoshida laboratories for help and advice; I. Mizuno, K. Suzuki, S. Uchiyama and Y. Mizuguchi for the photon-counting system; C. Gosse and H. Miyajima for the magnetic tweezers; K. Abe and K. Rikukawa for microscopy; S. Murakami for luciferase; and H. Umezawa and M. Fukatsu for laboratory management. This work was supported in part by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and Burroughs Wellcome Fund (R.Y.).

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to H.I. (hiritoh@hpk.tre-net.co.jp).

Contacts

Publisher: Ben Crowe
Editor: Paul Smaglik
Marketing Manager: David Bowen

European Head Office, London

The Macmillan Building
4 Crinan Street
London N1 9XW, UK
Tel +44 (0) 20 7843 4961
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs Sales Director:
Nevin Bayouni (4978)

UK/ RoW/ Ireland:

Matt Powell (4953)
Andy Douglas (4975)
Frank Phelan (4944)

Scandinavia/ Spain/ Portugal:

Evelina Rubio Håkansson (4973)

Natureevents: Silke Opstrup (4994)

France/ Switzerland:

Amelie Pequinot (4974)

Production Manager:

Billie Franklin
To send materials use London
address above.
Tel +44 (0) 20 7843 4814
Fax +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs web development:

Tom Hancock

Naturejobs online production:

Ben Lund

European Satellite Office

Germany/ Austria/ Netherlands/
Italy/ Belgium:

Patrick Phelan, Odo Wulffen
Tel + 49 89 54 90 57 11/-2
Fax + 49 89 54 90 57 20
e-mail: p.phelan@nature.com
o.wulffen@nature.com

US Head Office, New York

345 Park Avenue South,
10th Floor, New York, NY 10010-1707
Tel +1 800 989 7718
Fax +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager:

Peter Bless

US Advertising Coordinator:

Linda Adam

Japan Head Office, Tokyo

MG Ichigaya Building (5F),
19-1 Haraikatomachi,
Shinjuku-ku,
Tokyo 162-0841
Tel +81 3 3267 8751
Fax +81 3 3267 8746

Asia-Pacific Sales Director:

Rinoko Asami
e-mail: rasami@naturejpn.com

naturejobs

Mapping life-science skills

Cartographers commissioned by the Sector Skills Council for Science, Engineering and Manufacturing Technologies (SEMTA) to map the skills needed to work in Britain's bioscience and biotech sector found one broad area of the professional landscape lacking. And a specific set of missing skills within that region might provide the most opportunities for young scientists.

To complete the 'functional map', SEMTA last year surveyed 35 of Britain's drug-discovery firms to see what skills they need to operate, and whether the government should provide training in any particular area (see *Nature* 421, 871; 2003). This year, a panel of scientists from academia, the drug industry and the biotech sector combed the results of the survey and reported back.

The good news? The panel found plenty of aptitude in basic business functions such as sales, marketing and human resources within the bioscience companies. In fact, there was only one area that needed work, but it was the area on which the success of these science-based companies ultimately rests — applied research and development. The need is particularly acute in areas such as analytical chemistry, which companies use to create, then screen, molecules to fight disease. But the biggest single need identified by the panel was for specialists in information management — for instance, professionals who can set up, maintain and track libraries of biological compounds. This is a subset of bioinformatics, itself one of the most in-demand scientific skills globally over the past few years.

But identifying which skills are needed and devising ways to meet that demand is only the first step, the panel admitted. The next — and biggest — challenge is attracting students to disciplines such as chemistry, where interest has been falling steadily for years. Perhaps news that job opportunities exist in the sector can help to reverse the trend.

Paul Smaglik
Naturejobs editor



Contents

CAREER VIEW

Scientists & Societies

Encouraging junior faculty members

Graduate Journal

Tshaka Cunningham

Movers

Markus Wenk and
Sonia Davila

p470

WWW.NATUREJOBS.COM

Career centre
Information on the
scientific job market

FOCUS

SPOTLIGHT

RECRUITMENT

ANNOUNCEMENTS

EVENTS

GRADUATE JOURNAL

The long and the short of it

My thesis work at Rockefeller University, New York, is aimed at finding out how HIV carries its genetic information into the cells it infects. This requires me to set long- and short-term scientific goals. But to pursue those goals, I may have to make similar-style career decisions.

In the short term, I would like to uncover the identity and activity of the cellular proteins involved in HIV nuclear import. In the long term, I want to discover small molecules that can block this step in the viral replication cycle and so might be used as drugs to combat HIV.

It is no coincidence that my short-term goals are more in line with traditional academics and my long-term goals are more applied. I strongly believe in the power and potential of 'translational research' to bridge the gap between basic science and new medical treatments.

Whether or not I stay in academia, I'll be looking for a place that supports this type of disease-oriented research and that welcomes young investigators who like to explore new ideas. The importance of this kind of research has never been more apparent to me than during my time at Rockefeller, where the school insignia reads: "Science for the benefit of humankind". That is what I hope to do throughout my career. ■

Tshaka Cunningham is a fifth-year graduate student at Rockefeller University in New York.

SCIENTISTS & SOCIETIES

Junior Faculty at the Karolinska Institute

One of the toughest transitions that a scientist makes is from being a postdoc to becoming a junior faculty member. To ease this transition, the Karolinska Institute in Stockholm, Sweden, last year launched Junior Faculty, a programme targeted at junior researchers with a PhD. The aim is to provide the tools that will allow these scientists to succeed in their careers. To meet this goal, the programme offers mentoring, career counselling, courses and networking opportunities.

One of the big challenges is reaching the programme's target group — the Karolinska Institute's 28 departments are quite widely spread out, both physically and thematically. In addition, researchers in the target group are often reasonably 'invisible' as they have no

formal positions and so aren't always on the personnel registers.

The solution to this was to create an information network with an 'ambassador' at each department. The ambassadors will identify the target group within their department, spread information about Junior Faculty activities and meet with the other ambassadors to plan events and raise issues.

At the Center for Surgical Sciences, for example, ambassador Olav Rooyackers is arranging a series of workshops on funding issues. Another ambassador, Zarina Kabir at the Neurotec Department, has brought junior researchers together to talk about career paths. So far these discussions have resulted in an initiative to create a structured career plan for the younger scientists in the department.

The Junior Faculty programme has also contributed to the restart of a passive network at the Karolinska Institute, WISE — Women In Science. More than 40 women took part in the first meeting and the activities within the network will include scientific meetings, seminars and lobbying.

Networks can provide a forum for the discussion of common interests and provide support to researchers for their everyday work. They also act as a forum for the exchange of ideas. Regular dialogue between researchers from different parts of the university can form the basis for constructive new thinking and favourable personal development. It could even result in you meeting your next collaborator or someone to publish with. ■

► www.ki.se/juniorfaculty

Johanna Nilsson is coordinator for the Junior Faculty programme.

MOVERS

Markus Wenk, assistant professor, Depts of Biochemistry and Biological Sciences, National University of Singapore; Sonia Davila, postdoctoral fellow, Genome Institute of Singapore



Advancing in your scientific career is difficult enough, but securing a tenure-track position at the same time as your partner who is also a scientist is quite another. Markus Wenk and Sonia

Davila took a drastic approach to solving this problem — they moved halfway around the world from the United States to Singapore this year.

In 2002, Wenk realized that he needed to establish himself as an independent researcher — impossible as a non-tenure-track researcher at Yale University. "My driver was to become independent and to steer my work on lipid function towards a 'systems-biology' approach," he says.

So he looked to Asia, because he was aware that opportunities there tended to include a high-tech component — something he was interested in exploring in his research focus of membrane-lipid biology. He soon found an opportunity in Singapore, where the government has created a huge 'Biopolis' to house an expanded national research effort (see *Nature* **425**, 746–747; 2003).

When Wenk received an offer from the National University of Singapore (NUS), things got complicated. First, he got a counter-offer, including tenure-track status, from Yale. Second, his partner Davila, who Wenk had met at Yale, needed to find a job as well. He turned down Yale's offer and Davila secured a position at the Genome Institute of Singapore. Both admit that their sense of adventure played

a part in their decision to move.

Wenk and Davila say that they have noticed a different culture within their new research groups compared with the ones they left behind in the United States, where labs tend to be shrouded in secrecy and characterized by competition. "The lab concept here is not as closed as in the States," says Davila.

And neither feels isolated from the West. Wenk plans to keep an adjunct appointment with Yale and is also evaluating the possibility of creating a Master's programme in tropical-disease research at the University of Basel in Switzerland, the NUS and the Novartis Institute of Tropical Diseases in Singapore.

One other thing that Wenk and Davila are happy with is Singapore's climate — especially Davila, who finds the tropical temperatures more akin to her native Spain. ■

CV **Markus Wenk 2001–2004:** Associate research scientist, Yale University School of Medicine
1997–2001: Postdoctoral associate, cell-biology department at Yale University School of Medicine
1997: PhD biophysics, University of Basel, Switzerland

Sonia Davila 2001–2004: Postdoctoral associate, Yale University School of Medicine; PhD in genetics, University of Santiago de Compostela, Spain